GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:08:28 ; Search time 1183 Seconds
(without alignments)
1057.106 Million cell updates/sec

Title: US-10-697-802A-42
Perfect score: 22
Sequence: 1 gcgtgcttaacacatgcaagtc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	22	100.0	73	6	A32046	A32046 DNA probe (
2	22	100.0	73	6	A32064	A32064 DNA probe (
3	22	100.0	74	1	FSP16S1	X53208 Frankia spe
4	22	100.0	74	1	KA16S1	X53190 Kibdelospor
5	22	100.0	80	11	CS001913	CS001913 Sequence
6	22	100.0	97	3	AY711993	AY711993 Unculture
7	22	100.0	99	3	AY710568	AY710568 Unculture
8	22	100.0	105	1	AF051385	AF051385 Actinomad
9	22	100.0	108	1	AF051382	AF051382 Actinomad
10	22	100.0	108	1	AF051383	AF051383 Actinomad
11	22	100.0	111	1	AF051381	AF051381 Actinomad
12	22	100.0	117	1	AF051377	AF051377 Actinomad
13	22	100.0	117	1	AF051378	AF051378 Actinomad
14	22	100.0	118	1	AF051379	AF051379 Actinomad
15	22	100.0	118	1	AF051380	AF051380 Actinomad
16	22	100.0	130	3	AY642393	AY642393 Unculture
17	22	100.0	136	3	AY858488	AY858488 Unculture
c 18	22	100.0	136	3	AY858540	AY858540 Unculture
c 19	22	100.0	137	3	AY858490	AY858490 Unculture
20	22	100.0	157	3	AY858487	AY858487 Unculture
c 21	22	100.0	175	3	AY858494	AY858494 Unculture
22	22	100.0	183	3	AY710916	AY710916 Unculture
23	22	100.0	186	3	AY858485	AY858485 Unculture
24	22	100.0	207	3	AY710617	AY710617 Unculture
25	22	100.0	214	1	AY198382	AY198382 Purfural-
26	22	100.0	214	10	AB167458	AB167458 Sus scrof
27	22	100.0	218	3	AY038545	AY038545 Unculture
28	22	100.0	219	3	AY875910	AY875910 Unculture
29	22	100.0	220	1	S72448	S72448 16S rRNA [F
30	22	100.0	221	1	S72447	S72447 16S rRNA [F
31	22	100.0	221	1	AJ880355	AJ880355 Mycobacte
c 32	22	100.0	231	1	AY688937	AY688937 Streptomy
c 33	22	100.0	231	1	AY688938	AY688938 Streptomy
c 34	22	100.0	231	1	AY688939	AY688939 Streptomy
35	22	100.0	241	3	AY271777	AY271777 Unculture
36	22	100.0	242	3	AY006713	AY006713 Unculture
37	22	100.0	243	3	AY710732	AY710732 Unculture
38	22	100.0	247	3	AY710934	AY710934 Unculture
39	22	100.0	253	3	AF411234	AF411234 Unculture
40	22	100.0	256	3	AF114641	AF114641 Unculture
41	22	100.0	258	1	AY451332	AY451332 Arthrobac
42	22	100.0	259	3	AF376160	AF376160 Unculture
43	22	100.0	261	3	AY710709	AY710709 Unculture
44	22	100.0	264	3	AY239556	AY239556 Unculture
45	22	100.0	265	3	AF114664	AF114664 Unculture

ALIGNMENTS

RESULT 1
A32046
LOCUS A32046 73 bp DNA linear PAT 08-DEC-1995
DEFINITION DNA probe (M.bovis) from patent EP0395292.
ACCESSION A32046
VERSION A32046.1 GI:1249501
KEYWORDS .
SOURCE synthetic construct
ORGANISM synthetic construct
other sequences; artificial sequences.
REFERENCE 1 (bases 1 to 73)
AUTHORS Barry,T.G., Gannon,B.X. and Powell,R.
TITLE Generation of specific probes for target nucleotide sequences
JOURNAL Patent: EP 0395292-A 21 31-OCT-1990;
Barry, Thomas Gerard; Gannon, Bernard Francis Xavier; BIORESEARCH IRELAND; Powell, Richard; UNIVERSITY COLLEGE GALWAY; Barry, Thomas Gerard; Gannon, Bernard Francis Xavier; BIORESEARCH IRELAND; Powell, Richard; UNIVERSITY COLLEGE GALWAY; Barry, Thomas Gerard; Gannon, Bernard Francis Xavier; EOLAS (trading as BioResearch Ireland) - The Irish Science and Technology Agency; Powell, Richard; UNIVERSITY COLLEGE GALWAY
FEATURES Location/Qualifiers
source 1. .73
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 73;
Best Local Similarity 100.0%; Pred. No. 2.4e+04;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 21 GCGTGCTTAACACATGCAAGTC 42

RESULT 2
A32064
LOCUS A32064 73 bp DNA linear PAT 08-DEC-1995
DEFINITION DNA probe (M.bovis) from patent EP0395292.

```
ACCESSION   A32064
VERSION     A32064.1  GI:1249519
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1  (bases 1 to 73)
  AUTHORS   Barry,T.G., Gannon,B.X. and Powell,R.
  TITLE     Generation of specific probes for target nucleotide sequences
  JOURNAL   Patent: EP 0395292-A 39 31-OCT-1990;
            Barry, Thomas Gerard; Gannon, Bernard Francis Xavier; BIORESEARCH
            IRELAND; Powell, Richard; UNIVERSITY COLLEGE GALWAY; Barry, Thomas
            Gerard; Gannon, Bernard Francis Xavier; BIORESEARCH IRELAND;
            Powell, Richard; UNIVERSITY COLLEGE GALWAY; Barry, Thomas Gerard;
            Gannon, Bernard Francis Xavier; EOLAS (trading as BioResearch
            Ireland) - The Irish Science and Technology Agency; Powell,
            Richard; UNIVERSITY COLLEGE GALWAY
FEATURES             Location/Qualifiers
     source          1. .73
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 73;
  Best Local Similarity  100.0%;  Pred. No. 2.4e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         21 GCGTGCTTAACACATGCAAGTC 42


RESULT 3
FSP16S1
LOCUS       FSP16S1                    74 bp    DNA     linear   BCT 06-JUN-2003
DEFINITION  Frankia spec. strain Ag45/Mut15 partial 16S rRNA, part 1.
ACCESSION   X53208
VERSION     X53208.1  GI:43421
KEYWORDS    16S ribosomal RNA; ribosomal RNA.
SOURCE      Frankia sp.
  ORGANISM  Frankia sp.
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Frankineae; Frankiaceae; Frankia.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Hahn,D., Lechevalier,M.P., Fischer,A. and Stackebrandt,E.
  TITLE     Evidence for a close phylogenetic relationship between members of
            the genera Frankia, Geodermatophilus, and 'Blastococcus' and
            emendation of the family Frankiaceae
  JOURNAL   Syst. Appl. Microbiol. 11, 236-242 (1989)
REFERENCE   2  (bases 1 to 74)
  AUTHORS   Stackebrandt,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (09-MAY-1990) Stackebrandt E
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Frankia sp."
                     /mol_type="genomic DNA"
                     /strain="Ag45/Mut15"
                     /db_xref="taxon:1855"
     rRNA            1. .>74
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 1;  Length 74;
  Best Local Similarity  100.0%;  Pred. No. 2.3e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         43 GCGTGCTTAACACATGCAAGTC 64


RESULT 4
KA16S1
LOCUS       KA16S1                     74 bp    DNA     linear   BCT 06-JUN-2003
DEFINITION  Kibdelosporangium aridum 16S rRNA (part. 1).
ACCESSION   X53190
VERSION     X53190.1  GI:43770
KEYWORDS    16S ribosomal RNA; ribosomal RNA.
SOURCE      Kibdelosporangium aridum
  ORGANISM  Kibdelosporangium aridum
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Pseudonocardineae; Pseudonocardiaceae; Kibdelosporangium.
REFERENCE   1  (bases 1 to 74)
  AUTHORS   Bowen,T., Stackebrandt,E., Dorsch,M. and Embley,T.M.
  TITLE     The phylogeny of Amycolata autotrophica, Kibdelosporangium aridum
            and Saccharothrix australiensis
  JOURNAL   J. Gen. Microbiol. 135, 2529-2536 (1989)
REFERENCE   2  (bases 1 to 74)
  AUTHORS   Stackebrandt,E.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-APR-1990) Stackebrandt E
COMMENT     the genus Kibdelosporangium is proposed to be classified in the
            family Pseudonocardiaceae
            see X53191 for downstream 16S rRNA seq, a range of unknown length
            was not sequenced.
FEATURES             Location/Qualifiers
     source          1. .74
                     /organism="Kibdelosporangium aridum"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:2030"
     rRNA            1. .>74
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 1;  Length 74;
  Best Local Similarity  100.0%;  Pred. No. 2.3e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         43 GCGTGCTTAACACATGCAAGTC 64


RESULT 5
CS001913
LOCUS       CS001913                   80 bp    DNA     linear   SYN 01-FEB-2005
DEFINITION  Sequence 11 from Patent WO2004097369.
ACCESSION   CS001913
VERSION     CS001913.1  GI:58424130
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   van den Boom,D. and Boecker,S.
  TITLE     Fragmentation-based methods and systems for de novo sequencing
  JOURNAL   Patent: WO 2004097369-A 11-NOV-2004;
            Sequenom, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .80
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Synthetic
                     polynucleotide sequence"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 11;  Length 80;
  Best Local Similarity  100.0%;  Pred. No. 2.2e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         35 GCGTGCTTAACACATGCAAGTC 56
```

```
RESULT 6
AY711993
LOCUS       AY711993                97 bp    DNA     linear   ENV 12-OCT-2004
DEFINITION  Uncultured Piscirickettsiaceae bacterium clone SIMO-456 16S
            ribosomal RNA gene, partial sequence.
ACCESSION   AY711993
VERSION     AY711993.1  GI:53773468
KEYWORDS    ENV.
SOURCE      uncultured Piscirickettsiaceae bacterium
  ORGANISM  uncultured Piscirickettsiaceae bacterium
            Bacteria; Proteobacteria; Gammaproteobacteria; Thiotrichales;
            Piscirickettsiaceae; environmental samples.
REFERENCE   1  (bases 1 to 97)
  AUTHORS   Moran,M.A., Whitman,W.B. and Ye,W.
  TITLE     Diversity of salt marsh prokaryotes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 97)
  AUTHORS   Moran,M.A., Whitman,W.B. and Ye,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-AUG-2004) Department of Marine Sciences, University
            of Georgia, Athens, GA 30602, USA
FEATURES             Location/Qualifiers
     source          1. .97
                     /organism="uncultured Piscirickettsiaceae bacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="lon=81.2699W, lat=31.3929N; surface
                     water collected on Jul 19, 2001, Sapelo Island Microbial
                     Observatory Dean Creek Marsh sampling site"
                     /db_xref="taxon:191470"
                     /clone="SIMO-456"
                     /environmental_sample
                     /country="USA: Georgia, Sapelo Island"
     rRNA            <1. .>97
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 97;
  Best Local Similarity  100.0%;  Pred. No. 2e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db          8 GCGTGCTTAACACATGCAAGTC 29


RESULT 7
AY710568
LOCUS       AY710568                99 bp    DNA     linear   ENV 12-OCT-2004
DEFINITION  Uncultured proteobacterium clone SIMO-1128 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AY710568
VERSION     AY710568.1  GI:53772045
KEYWORDS    ENV.
SOURCE      uncultured proteobacterium
  ORGANISM  uncultured proteobacterium
            Bacteria; Proteobacteria; environmental samples.
REFERENCE   1  (bases 1 to 99)
  AUTHORS   Moran,M.A., Whitman,W.B. and Ye,W.
  TITLE     Diversity of salt marsh prokaryotes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 99)
  AUTHORS   Moran,M.A., Whitman,W.B. and Ye,W.
  TITLE     Direct Submission
  JOURNAL   Submitted (05-AUG-2004) Department of Marine Sciences, University
            of Georgia, Athens, GA 30602, USA
FEATURES             Location/Qualifiers
     source          1. .99
                     /organism="uncultured proteobacterium"
                     /mol_type="genomic DNA"
                     /isolation_source="lon=81.2797W, lat=31.3884N; sediment
                     0-1cm collected on Feb 01, 2002, Sapelo Island Microbial
                     Observatory Dean Creek Marsh sampling site"
                     /db_xref="taxon:153809"
                     /clone="SIMO-1128"
                     /environmental_sample
                     /country="USA: Georgia, Sapelo Island"
     rRNA            <1. .>99
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 99;
  Best Local Similarity  100.0%;  Pred. No. 2e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db          8 GCGTGCTTAACACATGCAAGTC 29


RESULT 8
AF051385
LOCUS       AF051385                105 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura viridis strain ATCC27103 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051385
VERSION     AF051385.1  GI:6652697
KEYWORDS    .
SOURCE      Actinomadura viridis
  ORGANISM  Actinomadura viridis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 105)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 105)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .105
                     /organism="Actinomadura viridis"
                     /mol_type="genomic DNA"
                     /strain="ATCC27103"
                     /db_xref="ATCC:27103"
                     /db_xref="taxon:58110"
     rRNA            <1. .>105
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 1;  Length 105;
  Best Local Similarity  100.0%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         13 GCGTGCTTAACACATGCAAGTC 34


RESULT 9
AF051382
LOCUS       AF051382                108 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura livida strain ATCC33578 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051382
VERSION     AF051382.1  GI:6652694
KEYWORDS    .
SOURCE      Actinomadura livida
  ORGANISM  Actinomadura livida
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 108)
```

```
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Actinomadura livida"
                     /mol_type="genomic DNA"
                     /strain="ATCC33578"
                     /db_xref="ATCC:33578"
                     /db_xref="taxon:79909"
     rRNA            <1. .>108
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         13 GCGTGCTTAACACATGCAAGTC 34


RESULT 10
AF051383
LOCUS       AF051383                108 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura viridis strain ATCC27888 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051383
VERSION     AF051383.1  GI:6652695
KEYWORDS    .
SOURCE      Actinomadura viridis
  ORGANISM  Actinomadura viridis
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 108)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 108)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .108
                     /organism="Actinomadura viridis"
                     /mol_type="genomic DNA"
                     /strain="ATCC27888"
                     /db_xref="ATCC:27888"
                     /db_xref="taxon:58110"
     rRNA            <1. .>108
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 108;
  Best Local Similarity   100.0%;  Pred. No. 1.9e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         13 GCGTGCTTAACACATGCAAGTC 34


RESULT 11
AF051381
LOCUS       AF051381                111 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura helvata strain ATCC27295 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051381
VERSION     AF051381.1  GI:6652693
KEYWORDS    .
SOURCE      Nonomuraea helvata
  ORGANISM  Nonomuraea helvata
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Streptosporangiaceae; Nonomuraea.
REFERENCE   1  (bases 1 to 111)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 111)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .111
                     /organism="Nonomuraea helvata"
                     /mol_type="genomic DNA"
                     /strain="ATCC27295"
                     /db_xref="ATCC:27295"
                     /db_xref="taxon:37484"
     rRNA            <1. .>111
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 111;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         14 GCGTGCTTAACACATGCAAGTC 35


RESULT 12
AF051377
LOCUS       AF051377                117 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura citrea strain ATCC27887 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051377
VERSION     AF051377.1  GI:6652689
KEYWORDS    .
SOURCE      Actinomadura citrea
  ORGANISM  Actinomadura citrea
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 117)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Actinomadura citrea"
                     /mol_type="genomic DNA"
                     /strain="ATCC27887"
                     /db_xref="ATCC:27887"
                     /db_xref="taxon:46158"
     rRNA            <1. .>117
                     /product="16S ribosomal RNA"
ORIGIN
```

```
  Query Match             100.0%;  Score 22;  DB 1;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         16 GCGTGCTTAACACATGCAAGTC 37


RESULT 13
AF051378
LOCUS       AF051378                 117 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura coerulea strain ATCC33576 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051378
VERSION     AF051378.1  GI:6652690
KEYWORDS    .
SOURCE      Actinomadura coerulea
  ORGANISM  Actinomadura coerulea
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 117)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 117)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .117
                     /organism="Actinomadura coerulea"
                     /mol_type="genomic DNA"
                     /strain="ATCC33576"
                     /db_xref="ATCC:33576"
                     /db_xref="taxon:46159"
     rRNA            <1. .>117
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 117;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         16 GCGTGCTTAACACATGCAAGTC 37


RESULT 14
AF051379
LOCUS       AF051379                 118 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura cremea subsp. cremea strain ATCC33577 16S ribosomal
            RNA gene, partial sequence.
ACCESSION   AF051379
VERSION     AF051379.1  GI:6652691
KEYWORDS    .
SOURCE      Actinomadura cremea subsp. cremea
  ORGANISM  Actinomadura cremea subsp. cremea
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 118)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .118
                     /organism="Actinomadura cremea subsp. cremea"
                     /mol_type="genomic DNA"
                     /strain="ATCC33577"
                     /db_xref="ATCC:33577"
                     /db_xref="taxon:31961"
     rRNA            <1. .>118
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         17 GCGTGCTTAACACATGCAAGTC 38


RESULT 15
AF051380
LOCUS       AF051380                 118 bp    DNA     linear   BCT 02-JAN-2000
DEFINITION  Actinomadura spadix strain ATCC27298 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   AF051380
VERSION     AF051380.1  GI:6652692
KEYWORDS    .
SOURCE      Actinomadura spadix
  ORGANISM  Actinomadura spadix
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Streptosporangineae; Thermomonosporaceae; Actinomadura.
REFERENCE   1  (bases 1 to 118)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Molecular Identification of Actinomycetes
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 118)
  AUTHORS   Rodriguez,V., Parro,V. and Mellado,R.P.
  TITLE     Direct Submission
  JOURNAL   Submitted (27-FEB-1998) Biotecnologia Microbiana, Centro Nacional
            de Biotecnologia, Campus de la Universidad Autonoma, Cantoblanco,
            Madrid 28049, Spain
FEATURES             Location/Qualifiers
     source          1. .118
                     /organism="Actinomadura spadix"
                     /mol_type="genomic DNA"
                     /strain="ATCC27298"
                     /db_xref="ATCC:27298"
                     /db_xref="taxon:79912"
     rRNA            <1. .>118
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 118;
  Best Local Similarity   100.0%;  Pred. No. 1.8e+04;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         17 GCGTGCTTAACACATGCAAGTC 38


Search completed: April  7, 2006, 20:42:17
Job time : 1187 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:01:48 ; Search time 220 Seconds
(without alignments)
666.469 Million cell updates/sec

Title: US-10-697-802A-42
Perfect score: 22
Sequence: 1 gcgtgcttaacacatgcaagtc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	% Query Match	Length	DB	ID	Description
1		22	100.0	22	14	AEA22441	Aea22441 Acid-fast
2		22	100.0	80	14	ADU66542	Adu66542 Cut base
3		22	100.0	166	2	AAX32481	Aax32481 Preferred
4		22	100.0	209	14	AEB98764	Aeb98764 Mycobacte
5		22	100.0	209	14	AEB98762	Aeb98762 Mycobacte
6		22	100.0	209	14	AEB98763	Aeb98763 Mycobacte
7		22	100.0	211	14	AEB98761	Aeb98761 Mycobacte
8		22	100.0	349	13	ADV99481	Adv99481 Meningiti
9		22	100.0	415	4	AAI92758	Aai92758 Human pol
10		22	100.0	421	2	AAV72337	Aav72337 Actinomyc
c 11		22	100.0	422	14	ADW94995	Adw94995 Clostridi
12		22	100.0	428	12	ADQ74829	Adq74829 Rhodococc
13		22	100.0	436	12	ADH48069	Adh48069 Arthrobac
14		22	100.0	447	12	ADQ74847	Adq74847 Rhodococc
15		22	100.0	460	8	ABZ76674	Abz76674 Microtetr
16		22	100.0	463	2	AAV72360	Aav72360 Actinomyc
17		22	100.0	463	3	AAZ57030	Aaz57030 Actinomyc
18		22	100.0	463	6	ABK88031	Abk88031 DNA encod
19		22	100.0	463	8	ABZ76675	Abz76675 Streptomy
20		22	100.0	463	8	ABZ76673	Abz76673 Streptomy
21		22	100.0	497	14	AEB72673	Aeb72673 Streptosp
22		22	100.0	500	13	ADZ20587	Adz20587 Formaldeh
23		22	100.0	500	14	AEA39586	Aea39586 Streptomy
24		22	100.0	500	14	AEB72672	Aeb72672 Streptosp
25		22	100.0	500	14	AEB98339	Aeb98339 16SrDNA s
26		22	100.0	501	12	ADP03611	Adp03611 DNA seque
27		22	100.0	502	10	ADF86316	Adf86316 Amycolato
28		22	100.0	502	12	ADP88197	Adp88197 Antagonis
29		22	100.0	502	12	ADP88198	Adp88198 Antagonis
30		22	100.0	503	12	ADP03610	Adp03610 DNA seque
31		22	100.0	503	12	ADP88195	Adp88195 Antagonis
32		22	100.0	503	12	ADP88196	Adp88196 Antagonis
33		22	100.0	535	13	ADS75567	Ads75567 Rhodococc
34		22	100.0	560	10	ABT23572	Abt23572 Stabilisi
35		22	100.0	560	10	ABT23571	Abt23571 Stabilisi
36		22	100.0	582	8	ACD26614	Acd26614 Puromycin
c 37		22	100.0	711	14	ADW16274	Adw16274 DNA copy
38		22	100.0	787	2	AAV43262	Aav43262 Partial 1
39		22	100.0	1312	4	AAF28889	Aaf28889 Arthrobac
40		22	100.0	1315	4	AAF28890	Aaf28890 Arthrobac
c 41		22	100.0	1343	12	ADO80217	Ado80217 Rhodococc
42		22	100.0	1343	14	AEA00984	Aea00984 16S ribos
c 43		22	100.0	1344	12	ADO85868	Ado85868 Gordonia
44		22	100.0	1388	10	ADC61230	Adc61230 Baeyer-Vi
45		22	100.0	1391	2	AAT45276	Aat45276 Corynebac

ALIGNMENTS

RESULT 1
AEA22441
ID AEA22441 standard; DNA; 22 BP.
XX
AC AEA22441;
XX
DT 25-AUG-2005 (first entry)
XX
DE Acid-fast bacterium forward (AFB-f) 16S rDNA PCR primer SEQ ID NO:42.
XX
KW microorganism identification; 16S rDNA; 16S ribosomal DNA; PCR; primer;
KW ss.
XX
OS Synthetic.
XX
PN US2005130168-A1.
XX
PD 16-JUN-2005.
XX
PF 31-OCT-2003; 2003US-00697802.
XX
PR 31-OCT-2003; 2003US-00697802.
XX
PA (HANX/) HAN X.
PA (PHAM/) PHAM A S.
XX
PI Han X, Pham AS;
XX
DR WPI; 2005-424597/43.
XX
PT Determining a bacterium species comprises providing oligonucleotide
PT primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS Claim 2; SEQ ID NO 42; 74pp; English.
XX
CC The invention relates to a method (M1) for determining a bacterium
CC species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC extracting a genomic nucleotide from the bacterium to provide a
CC nucleotide template; (c) annealing a region of a nucleotide template to a
CC specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC complimentary fashion, the primer set designed to provide a product
CC having a predetermined size dictated by a complimentary primer set; (d)

CC amplifying the region of the nucleotide template to produce the product;
CC and (e) determining a species of a bacterium in a nucleotide sequence of
CC the product. Also described is an alternative method (M2) for determining
CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or
CC variations, or (iii) a first forward primer having consecutive bases of
CC an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC second reverse primer having consecutive bases of an UB-r of AEA22517-
CC AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC comparing the product from the specimen with a nucleotide sequence from a
CC database to determine the bacterium species present in the specimen. The
CC methods are useful for determining a bacterium species. The present
CC sequence represents a forward PCR primer for amplifying 16S rDNA regions
CC of acid-fast bacterium (AFB), which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 22 BP; 6 A; 6 C; 5 G; 5 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.32;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 1 GCGTGCTTAACACATGCAAGTC 22

RESULT 2
ADU66542
ID ADU66542 standard; DNA; 80 BP.
XX
AC ADU66542;
XX
DT 27-JAN-2005 (first entry)
XX
DE Cut base A amplicon fragment.
XX
KW ds; mass spectroscopy; DNA cleavage; DNA sequencing; sequencing.
XX
OS Unidentified.
XX
PN WO2004097369-A2.
XX
PD 11-NOV-2004.
XX
PF 22-APR-2004; 2004WO-US012520.
XX
PR 25-APR-2003; 2003US-0466006P.
XX
PA (SEQU-) SEQUENOM INC.
PA (BOEC/) BOECKER S.
XX
PI Boecker S, Van Den Boom D;
XX
DR WPI; 2005-012656/01.
XX
PT Obtaining sequence information from target biomolecule, by fragmenting
PT target biomolecule by partial cleavage, performing mass spectrometry,
PT extracting information from mass spectra, constructing sequencing graph
PT and traversing graphs.
XX
PS Disclosure; SEQ ID NO 11; 133pp; English.
XX
CC This invention describes a novel method for obtaining sequence
CC information from a target biomolecule and involves fragmenting the target
CC biomolecule into several fragments by partial cleavage, performing mass
CC spectrometry on fragments to produce mass spectra, extracting peak
CC information from the produced mass spectra, constructing sequencing graph
CC using the extracted peak information and traversing the sequencing graphs
CC to reconstruct sequence information of the target biomolecule. The target
CC biomolecule is nucleic acid molecule such as DNA or RNA, or is a protein
CC and the compositions of the two fragments are the base compositions or
CC amino acid compositions. This method preferably involves subjecting the
CC nucleic acid molecule to partial cleavage reactions with one or more
CC specific cleavage reagents, thus generating two or more fragments that
CC are specific cleavage products, determining the molecular weights of the
CC two or more fragments, determining the possible base compositions of the
CC two or more fragments, ordering the possible base compositions of the two
CC or more fragments according to the number of specific cleavage sites that
CC are not cleaved in each fragment, constructing one or more sequencing
CC graphs that are a graph theoretical representation of the ordered base
CC compositions for the two or more fragments, and traversing the one or
CC more sequencing graph to reconstruct one or more underlying sequence
CC candidates, where each sequencing graph corresponds to the ordered base
CC compositions derived from a partial cleavage reaction with one base-
CC specific cleavage reagent. This method further involves scoring the one
CC or more underlying sequence candidates and determining the rank order of
CC fitness, where the scoring is done by statistical analysis or maximum
CC likelihood statistical analysis. This method determines epigenetic
CC changes in a target nucleic acid molecule relative to reference nucleic
CC acid molecule and allows the sequencing of large biomolecules. The
CC invention also describes a method of producing a candidate sequence of a
CC biomolecule which involves receiving several sequencing graphs having
CC several vertices and edges, where each vertex represents a compomer of
CC the biomolecule and each edge represents a cut base of the sequencing
CC graph and generating the candidate sequence by traversing several
CC sequencing graphs. This second method further involves traversing several
CC sequencing graphs by tracing through each sequencing graph, starting at a
CC source vertex. The results of each method can be read by a program
CC product for use in a computer that executes program instructions recorded
CC in a computer-readable media to produce a candidate sequence of a
CC biomolecule or to obtain sequence information in a target biomolecule.
CC The target biomolecule contains a sequence variation, which is a mutation
CC or a polymorphism. The target is a target nucleic acid molecule from an
CC organism chosen from eukaryotes, prokaryotes and viruses, preferably a
CC bacterium. The specific cleavage reagent is an RNAse chosen from RNase
CC T1, RNase U2, RNase PhyM, RNase A, chicken liver RNase (RNase CL3) and
CC cusavitin, or a glycosylase. The sequence variations in the target
CC biomolecule permit genotyping a subject, forensic analysis, disease
CC diagnosis or disease prognosis. The novel methods are useful for de novo
CC sequencing, to identify genetic disease or chromosome abnormality,
CC identifying a predisposition to a disease, or condition including
CC obesity, atherosclerosis, or cancer, to identify an infection by an
CC infectious agent, to identify a pathogens, determine haplotypes, analyze
CC microsatellite sequences, and short tandem repeat (STR) loci, determine
CC allelic variation and/or frequency, and analyze cellular methylation
CC patterns. This sequence represents an amplicon used to illustrate the
CC sequencing technique described in the invention.
XX
SQ Sequence 80 BP; 18 A; 20 C; 27 G; 15 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 80;
Best Local Similarity 100.0%; Pred. No. 0.38;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 35 GCGTGCTTAACACATGCAAGTC 56

RESULT 3
AAX32481
ID AAX32481 standard; DNA; 166 BP.
XX
AC AAX32481;
XX
DT 22-JUN-1999 (first entry)

XX
DE Preferred probe of the invention.
XX
KW 16S rRNA; maduromycetes; hybridisation; streptomycetes; probe; ss.
XX
OS Synthetic.
OS Streptomyces ambofaciens.
XX
PN WO9914361-A1.
XX
PD 25-MAR-1999.
XX
PF 16-SEP-1998; 98WO-EP006038.
XX
PR 18-SEP-1997; 97US-0059295P.
PR 16-DEC-1997; 97US-0069748P.
XX
PA (MERI) MERCK SHARP & DOHME ESPANA SAE.
XX
PI Genilloud O, Mellado RP, Parro V, Rodriguez V;
XX
DR WPI; 1999-229548/19.
XX
PT New probes used for detection of maduromycetes bacteria and to
PT differentiate between maduromycetes and streptomycetes.
XX
PS Disclosure; Fig 1; 22pp; English.
XX
CC The invention relates to a novel nucleic acid probe hybridises to a
CC nucleic acid encoding a portion of 16S rRNA of maduromycetes bacteria
CC under hybridisation conditions, and does not hybridise to nucleic acids
CC encoding a portion of 16S rRNA of streptomycetes under identical
CC hybridisation conditions. The probes can be used for detecting the
CC presence of maduromycetes bacteria in a sample and for differentiating
CC between maduromycetes and streptomycetes bacteria in a sample. The
CC present sequence represents a preferred probe of the invention
XX
SQ Sequence 166 BP; 39 A; 43 C; 55 G; 29 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 2; Length 166;
Best Local Similarity 100.0%; Pred. No. 0.42;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 41 GCGTGCTTAACACATGCAAGTC 62

RESULT 4
AEB98764
ID AEB98764 standard; DNA; 209 BP.
XX
AC AEB98764;
XX
DT 06-OCT-2005 (first entry)
XX
DE Mycobacterium kansasii partial 16S rDNA sequence, SEQ ID 6.
XX
KW microorganism detection; mycobacterium infection; antibacterial; ds.
XX
OS Mycobacterium kansasii.
XX
PN JP2005204582-A.
XX
PD 04-AUG-2005.
XX
PF 23-JAN-2004; 2004JP-00015195.
XX
PR 23-JAN-2004; 2004JP-00015195.
XX
PA (ASAH) ASAHI KASEI KK.
XX
PI Oda N;
XX
DR WPI; 2005-526965/54.
XX
PT New single-stranded oligonucleotide, useful for amplifying the nucleic
PT acid of Mycobacterium avium, Mycobacterium intracellulare, and
PT Mycobacterium kansasii.
XX
PS Example 1; SEQ ID NO 6; 14pp; Japanese.
XX
CC The invention relates to a novel single-stranded oligonucleotide used in
CC a detection method of an atypical mycobacteria group. The invention
CC further includes: amplifying the nucleic acid of Mycobacterium avium by a
CC loop-mediated isothermal amplification (LAMP) method; amplifying the
CC nucleic acid of M. intracellulare by a LAMP method; amplifying the
CC nucleic acid of M. kansasii by a LAMP method; and a kit for detecting the
CC nucleic acid of M. avium by a LAMP method; detecting the nucleic acid of
CC M. intracellulare by a LAMP method; or detecting the nucleic acid of M.
CC kansasii by a LAMP method. The single-stranded oligonucleotide is useful
CC in medical applications. This polynucleotide represents a Mycobacterium
CC kansasii partial 16S rDNA sequence amplified by the LAMP method of the
CC invention.
XX
SQ Sequence 209 BP; 47 A; 49 C; 72 G; 41 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 24 GCGTGCTTAACACATGCAAGTC 45

RESULT 5
AEB98762
ID AEB98762 standard; DNA; 209 BP.
XX
AC AEB98762;
XX
DT 06-OCT-2005 (first entry)
XX
DE Mycobacterium avium partial 16S rDNA sequence, SEQ ID 4.
XX
KW microorganism detection; mycobacterium infection; antibacterial; ds.
XX
OS Mycobacterium avium.
XX
PN JP2005204582-A.
XX
PD 04-AUG-2005.
XX
PF 23-JAN-2004; 2004JP-00015195.
XX
PR 23-JAN-2004; 2004JP-00015195.
XX
PA (ASAH) ASAHI KASEI KK.
XX
PI Oda N;
XX
DR WPI; 2005-526965/54.
XX
PT New single-stranded oligonucleotide, useful for amplifying the nucleic
PT acid of Mycobacterium avium, Mycobacterium intracellulare, and
PT Mycobacterium kansasii.
XX
PS Example 1; SEQ ID NO 4; 14pp; Japanese.
XX
CC The invention relates to a novel single-stranded oligonucleotide used in
CC a detection method of an atypical mycobacteria group. The invention
CC further includes: amplifying the nucleic acid of Mycobacterium avium by a
CC loop-mediated isothermal amplification (LAMP) method; amplifying the
CC nucleic acid of M. intracellulare by a LAMP method; amplifying the
CC nucleic acid of M. kansasii by a LAMP method; and a kit for detecting the

CC nucleic acid of M. avium by a LAMP method; detecting the nucleic acid of
CC M. intracellulare by a LAMP method; or detecting the nucleic acid of M.
CC kansasii by a LAMP method. The single-stranded oligonucleotide is useful
CC in medical applications. This polynucleotide represents a Mycobacterium
CC avium partial 16S rDNA sequence amplified by the LAMP method of the
CC invention.
XX
SQ Sequence 209 BP; 48 A; 48 C; 70 G; 43 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 24 GCGTGCTTAACACATGCAAGTC 45

RESULT 6
AEB98763
ID AEB98763 standard; DNA; 209 BP.
XX
AC AEB98763;
XX
DT 06-OCT-2005 (first entry)
XX
DE Mycobacterium intracellulare partial 16S rDNA sequence, SEQ ID 5.
XX
KW microorganism detection; mycobacterium infection; antibacterial; ds.
XX
OS Mycobacterium intracellulare.
XX
PN JP2005204582-A.
XX
PD 04-AUG-2005.
XX
PF 23-JAN-2004; 2004JP-00015195.
XX
PR 23-JAN-2004; 2004JP-00015195.
XX
PA (ASAH) ASAHI KASEI KK.
XX
PI Oda N;
XX
DR WPI; 2005-526965/54.
XX
PT New single-stranded oligonucleotide, useful for amplifying the nucleic
PT acid of Mycobacterium avium, Mycobacterium intracellulare, and
PT Mycobacterium kansasii.
XX
PS Example 1; SEQ ID NO 5; 14pp; Japanese.
XX
CC The invention relates to a novel single-stranded oligonucleotide used in
CC a detection method of an atypical mycobacteria group. The invention
CC further includes: amplifying the nucleic acid of Mycobacterium avium by a
CC loop-mediated isothermal amplification (LAMP) method; amplifying the
CC nucleic acid of M. intracellulare by a LAMP method; amplifying the
CC nucleic acid of M. kansasii by a LAMP method; and a kit for detecting the
CC nucleic acid of M. avium by a LAMP method; detecting the nucleic acid of
CC M. intracellulare by a LAMP method; or detecting the nucleic acid of M.
CC kansasii by a LAMP method. The single-stranded oligonucleotide is useful
CC in medical applications. This polynucleotide represents a Mycobacterium
CC intracellulare partial 16S rDNA sequence amplified by the LAMP method of
CC the invention.
XX
SQ Sequence 209 BP; 45 A; 47 C; 73 G; 44 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 209;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 24 GCGTGCTTAACACATGCAAGTC 45

RESULT 7
AEB98761
ID AEB98761 standard; DNA; 211 BP.
XX
AC AEB98761;
XX
DT 06-OCT-2005 (first entry)
XX
DE Mycobacterium tuberculosis partial 16S rDNA sequence, SEQ ID 3.
XX
KW microorganism detection; mycobacterium infection; antibacterial; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN JP2005204582-A.
XX
PD 04-AUG-2005.
XX
PF 23-JAN-2004; 2004JP-00015195.
XX
PR 23-JAN-2004; 2004JP-00015195.
XX
PA (ASAH) ASAHI KASEI KK.
XX
PI Oda N;
XX
DR WPI; 2005-526965/54.
XX
PT New single-stranded oligonucleotide, useful for amplifying the nucleic
PT acid of Mycobacterium avium, Mycobacterium intracellulare, and
PT Mycobacterium kansasii.
XX
PS Example 1; SEQ ID NO 3; 14pp; Japanese.
XX
CC The invention relates to a novel single-stranded oligonucleotide used in
CC a detection method of an atypical mycobacteria group. The invention
CC further includes: amplifying the nucleic acid of Mycobacterium avium by a
CC loop-mediated isothermal amplification (LAMP) method; amplifying the
CC nucleic acid of M. intracellulare by a LAMP method; amplifying the
CC nucleic acid of M. kansasii by a LAMP method; and a kit for detecting the
CC nucleic acid of M. avium by a LAMP method; detecting the nucleic acid of
CC M. intracellulare by a LAMP method; or detecting the nucleic acid of M.
CC kansasii by a LAMP method. The single-stranded oligonucleotide is useful
CC in medical applications. This polynucleotide represents a Mycobacterium
CC tuberculosis partial 16S rDNA sequence amplified by the LAMP method of
CC the invention.
XX
SQ Sequence 211 BP; 48 A; 45 C; 74 G; 44 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 211;
Best Local Similarity 100.0%; Pred. No. 0.44;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 24 GCGTGCTTAACACATGCAAGTC 45

RESULT 8
ADV99481
ID ADV99481 standard; DNA; 349 BP.
XX
AC ADV99481;
XX
DT 24-FEB-2005 (first entry)
XX
DE Meningitis causing bacteria DNA fragment #9.
XX
KW ds; antibacterial; antiinflammatory; inflammation; neurological disease;
KW diagnosis; meningitis; biochip.

XX
OS Mycobacterium tuberculosis.
XX
PN CN1420123-A.
XX
PD 28-MAY-2003.
XX
PF 16-NOV-2001; 2001CN-00137478.
XX
PR 16-NOV-2001; 2001CN-00137478.
XX
PA (JING-) JINGQI BIO CHEM SCI & TECH CO LTD.
XX
PI Xu B, Jiang Y, Huang X;
XX
DR WPI; 2004-044307/05.
XX
PT A nucleic acid sequence useful for diagnosing pathogenic bacteria for
PT meningitides.
XX
PS Disclosure; Page 18; 24pp; Chinese.
XX
CC The invention relates to a nucleic acid sequence group for quickly
CC diagnosing 20 kinds of pathogenic bacteria for meningitis. Its method
CC includes comparing the DNA sequences of different pathogenic bacteria,
CC choosing special fragments, finding out common primer, designing 3
CC specific probe fragments for each pathogenic bacterium, dotting them on
CC high-molecular polymer to obtain chip, sampling the DNA of pathogenic
CC bacterium of patient, labeling, amplification, and reacting with said
CC chip for visully recognizing the pathogenic bacterium. Its advantages are
CC high speed and low cost. The present sequence represents a meningitis
CC causing bacteria DNA fragment.
XX
SQ Sequence 349 BP; 75 A; 82 C; 125 G; 67 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 13; Length 349;
Best Local Similarity 100.0%; Pred. No. 0.47;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 37 GCGTGCTTAACACATGCAAGTC 58

RESULT 9
AAI92758
ID AAI92758 standard; cDNA; 415 BP.
XX
AC AAI92758;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 12818.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO12827.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 12818; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 415 BP; 86 A; 108 C; 145 G; 76 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 4; Length 415;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 47 GCGTGCTTAACACATGCAAGTC 68

RESULT 10
AAV72337
ID AAV72337 standard; DNA; 421 BP.
XX
AC AAV72337;
XX
DT 27-AUG-2003 (revised)
DT 28-JUL-1999 (first entry)
XX
DE Actinomycete sp. 16S rRNA DNA.
XX
KW Cellulase; detergent; animal feed; nutritional value; textile;
KW stone washing; texture modification; appearance; cellulosic fabric; pulp;
KW draining; paper; baking additive; starch treatment; grain;
KW high-fructose corn syrup production; ethanol production; fibre reduction;
KW milling; 16S rRNA; ss.
XX
OS Actinomyces sp.
XX
PN WO9925847-A2.
XX
PD 27-MAY-1999.
XX
PF 18-NOV-1998; 98WO-US024650.
XX
PR 19-NOV-1997; 97US-00974041.
PR 19-NOV-1997; 97US-00974042.
PR 22-JUN-1998; 98US-00102204.
XX
PA (GEMV) GENENCOR INT INC.
XX
PI Jones BE, Van Der Kleij WAH, Van Solingen P, Weyler W;
XX
DR WPI; 1999-347482/29.
XX
PT Cellulase from Actinomycetes.
XX
PS Example 4; Fig 6; 37pp; English.
XX

```
CC   This invention describes a novel cellulase isolated from an Actinomycete
CC   sp. which can be used in detergent compositions, as animal feeds (to
CC   increase nutritional value) and in treatment of textiles (e.g. stone
CC   washing or modifying texture, feel and/or appearance of cellulosic
CC   fabrics, including removal of 'immature' or 'dead' cotton), pulp (to
CC   improve draining) and paper. They may also be used as baking additives,
CC   for treating starch (in production of high-fructose corn syrup or
CC   ethanol) and for treating grain (to reduce fibre during milling).
CC   (Updated on 27-AUG-2003 to correct OS field.)
XX
SQ   Sequence 421 BP; 93 A; 108 C; 146 G; 74 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 2;  Length 421;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         28 GCGTGCTTAACACATGCAAGTC 49


RESULT 11
ADW94995/c
ID   ADW94995 standard; DNA; 422 BP.
XX
AC   ADW94995;
XX
DT   21-APR-2005  (first entry)
XX
DE   Clostridium botulinum 16S ribosomal RNA gene fragment, SEQ ID 2.
XX
KW   Antibacterial; Gastrointestinal-Gen.; Vaccine; microorganism;
KW   16s ribosomal RNA; 16s rRNA; enteropathy; ds.
XX
OS   Clostridium botulinum.
XX
PN   FR2858330-A1.
XX
PD   04-FEB-2005.
XX
PF   01-AUG-2003; 2003FR-00009562.
XX
PR   01-AUG-2003; 2003FR-00009562.
XX
PA   (CEVA-) CEVA SANTE ANIMALE SA.
XX
PI   Butty PJL;
XX
DR   WPI; 2005-134516/15.
DR   GENBANK; L37588.
XX
PT   New species Clostridium buttylinum, useful in vaccines for treatment and
PT   prevention of enteropathy in rabbits.
XX
PS   Disclosure; SEQ ID NO 2; 52pp; French.
XX
CC   The present invention relates to a novel species of bacterium,
CC   Clostridium buttylinum (Cb), which was isolated from rabbits. Cb is
CC   phytogenetically close to C. botulinum, C. novyi, C. sporogenes and C.
CC   sordelli. The first 420 nucleotides of its 16S ribosomal RNA gene is over
CC   95% identical with ADW94994. Cb, or compositions containing it, are used
CC   to prepare vaccines for prevention and/or treatment of enteropathy in
CC   rabbits, particularly rabbit epizootic and/or mucoid enteropathies. Cb
CC   was deposited with the Collection Nationale de Cultures de
CC   Microorganismes (CNCM) under number CNCM I-3029. The present sequence was
CC   used in a sequence homology alignment with the ADW94994 sequence of Cb.
XX
SQ   Sequence 422 BP; 89 A; 132 C; 91 G; 110 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 14;  Length 422;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        421 GCGTGCTTAACACATGCAAGTC 400


RESULT 12
ADQ74829
ID   ADQ74829 standard; DNA; 428 BP.
XX
AC   ADQ74829;
XX
DT   09-SEP-2004  (first entry)
XX
DE   Rhodococcus pyridinivorans dioxin associated 16s rDNA.
XX
KW   Rhodococcus sp. Probio-43; dioxin; dioxin-degrading activity; wastewater;
KW   sewage; river; sea; soil; 16s rDNA; ds.
XX
OS   Rhodococcus pyridinivorans.
XX
PN   KR2003091605-A.
XX
PD   03-DEC-2003.
XX
PF   28-MAY-2002; 2002KR-00029721.
XX
PR   28-MAY-2002; 2002KR-00029721.
XX
PA   (PROB-) PROBIONIC INC.
XX
PI   Cho YG,  Lee IS,  Park YH,  Yoon JH;
XX
DR   WPI; 2004-264438/25.
XX
PT   Novel microorganism Rhodococcus sp. probio-43 degrading dioxin.
XX
PS   Example 2; SEQ ID NO 1; 6pp; Korean.
XX
CC   The invention describes a novel microorganism Rhodococcus sp. Probio-43
CC   degrading dioxin, which effectively degrades and removes dioxin from the
CC   environment. A novel microorganism Rhodococcus sp. Probio-43 (KCCM 10380)
CC   is characterized by having dioxin-degrading activity. Also described is a
CC   composition for removing dioxin from wastewater, sewage, river, sea or
CC   soil characteristically contains Rhodococcus sp. Probio-43 (KCCM 10380).
CC   This sequence represents Rhodococcus pyridinivorans 16s rDNA associated
CC   with the degrading dioxin of the invention.
XX
SQ   Sequence 428 BP; 94 A; 105 C; 149 G; 80 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 12;  Length 428;
  Best Local Similarity   100.0%;  Pred. No. 0.48;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         15 GCGTGCTTAACACATGCAAGTC 36


RESULT 13
ADH48069
ID   ADH48069 standard; DNA; 436 BP.
XX
AC   ADH48069;
XX
DT   25-MAR-2004  (first entry)
XX
DE   Arthrobacter nicotianae 16S rDNA sequence SEQ ID NO:3.
XX
KW   alpha-H-alpha-amino acid amide racemase; enzyme; microorganism;
KW   racemisation; enantiomerically enriched alpha-H-alpha-amino acid amide;
KW   L-alpha-H-alpha-amino acid; D-alpha-H-alpha-amino acid amide;
KW   D-alpha-H-alpha-amino acid; L-alpha-H-alpha-amino acid amide;
KW   enantioselective amidase; gene; ds; 16S rDNA.
```

XX
OS Arthrobacter nicotianae.
XX
PN WO2003106691-A1.
XX
PD 24-DEC-2003.
XX
PF 13-JUN-2003; 2003WO-NL000423.
XX
PR 14-JUN-2002; 2002EP-00100711.
PR 20-DEC-2002; 2002EP-00080631.
XX
PA (STAM) DSM IP ASSETS BV.
XX
PI Boesten WHJ, Raemakers-Franken PC, Sonke T, Euverink GJW;
PI Grijpstra P;
XX
DR WPI; 2004-099017/10.
XX
PT Novel isolated Ochrobactrum anthropi 1A or Arthrobacter nicotianae alpha-
PT H-alpha-amino acid amide racemase polypeptide, useful for racemization of
PT an enantiomerically enriched alpha-H-alpha-amino acid amide.
XX
PS Example 2; SEQ ID NO 3; 78pp; English.
XX
CC The present invention describes an alpha-H-alpha-amino acid amide
CC racemase (I). Also described: (1) isolated fusion protein (II) made by
CC expression of a nucleic acid sequence encoding (I) operatively linked to
CC one or more nucleic acid sequences, which encode (a) marker
CC polypeptide(s); (2) nucleic acid sequence (III) encoding (I) or (II); (3)
CC vector (IV) comprising (III); (4) host cell (V) comprising and expressing
CC (III) or (IV); (5) isolating (M1) a microorganism displaying alpha-H-
CC alpha-amino acid amide racemase activity; (6) microorganism (VI)
CC obtainable by (M1); (7) Agrobacterium rhizogenes Na deposited under
CC number NCIMB 41127, A. rhizogenes Bi deposited under number NCIMB 41128,
CC Arthrobacter nicotianae deposited under number NCIMB 41126, Ochrobactrum
CC anthropi IA deposited under number NCIMB 41129; (8) isolating (M2) a
CC nucleic acid encoding polypeptide with alpha-H-alpha-amino acid amide
CC racemase activity, involves carrying out (M1), and isolating the nucleic
CC acid sequence from the obtained microorganism(s) by a standard method;
CC (9) nucleic acid sequence (VII) obtainable by (M2); (10) preparation of
CC (I); and (11) polypeptide produced by above mentioned method. (I), (V) or
CC (VI) can be used for racemisation of an enantiomerically enriched alpha-H
CC -alpha-amino acid amide, where the racemisation is performed in the
CC presence of (I), in the presence of (V) or (VI). (I), (V) or (VI) can
CC also be used for preparing enantiomerically enriched alpha-H-alpha-amino
CC acid amides or for preparing L-alpha-H-alpha-amino acid from the
CC corresponding D-alpha-H-alpha-amino acid amide or for preparing D-alpha-H
CC -alpha-amino acid from the corresponding L-alpha-H-alpha-amino acid
CC amide, which involves carrying out the process in the presence of an
CC enantioselective amidase and in the presence of (I), (V) or (VI). The
CC present sequence represents the 16S rDNA sequence of Arthrobacter
CC nicotianae NCIMB 41126, which is used in an example from the present
CC invention.
XX
SQ Sequence 436 BP; 97 A; 107 C; 151 G; 81 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 12; Length 436;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 3 GCGTGCTTAACACATGCAAGTC 24

RESULT 14
ADQ74847
ID ADQ74847 standard; DNA; 447 BP.
XX
AC ADQ74847;
XX
DT 09-SEP-2004 (first entry)
XX
DE Rhodococcus zopfii dioxin degradation associated 16s rDNA.
XX
KW Rhodococcus sp. Probio-42; dioxin; dioxin-degrading activity; wastewater;
KW sewage; river; sea; soil; 16s rDNA; ds.
XX
OS Rhodococcus zopfii.
XX
PN KR2003091604-A.
XX
PD 03-DEC-2003.
XX
PF 28-MAY-2002; 2002KR-00029720.
XX
PR 28-MAY-2002; 2002KR-00029720.
XX
PA (PROB-) PROBIONIC INC.
XX
PI Cho YG, Lee IS, Park YH, Yoon JH;
XX
DR WPI; 2004-278632/26.
XX
PT Novel microorganism rhodococcus sp. probio-42 degrading dioxin.
XX
PS Example 2; SEQ ID NO 1; 6pp; Korean.
XX
CC The invention describes a novel microorganism Rhodococcus sp. Probio-42
CC degrading dioxin, which effectively degrades and removes dioxin from the
CC environment. A novel microorganism Rhodococcus sp. Probio-42 (KCCM 10379)
CC is characterised by having dioxin-degrading activity. Also described is a
CC composition for removing dioxin from wastewater, sewage, river, sea or
CC soil characteristically contains Rhodococcus sp. Probio-42 (KCCM 10379).
CC This sequence represents Rhodococcus zopfii 16s rDNA associated with
CC dioxin degradation.
XX
SQ Sequence 447 BP; 103 A; 106 C; 158 G; 80 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 12; Length 447;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 15 GCGTGCTTAACACATGCAAGTC 36

RESULT 15
ABZ76674
ID ABZ76674 standard; DNA; 460 BP.
XX
AC ABZ76674;
XX
DT 30-APR-2003 (first entry)
XX
DE Microtetraspora recticatena IFO14525 DNA sequence SEQ ID NO:5.
XX
KW Streptomyces sp. TM-7; pravastatin; compactin; hyperlipidaemia;
KW antilipaemic; microorganism; gene; ds.
XX
OS Nonomuraea recticatena.
XX
PN WO200299109-A1.
XX
PD 12-DEC-2002.
XX
PF 30-MAY-2002; 2002WO-JP005252.
XX
PR 01-JUN-2001; 2001JP-00166412.
XX
PA (SAOC) MERCIAN CORP.
XX
PI Fujii T, Hirosue S, Aritoku Y, Morimiya T, Johdo O, Isshiki K;
XX

DR WPI; 2003-148672/14.
XX
PT Novel Streptomyces sp. produced polypeptide for hydroxylation of
PT compactin at 6beta-position and its encoded DNA, applicable in
PT constructing transformant microbes to synthesize pravastatin for treating
PT hyperlipedemia.
XX
PS Disclosure; Page 50-51; 67pp; Japanese.
XX
CC The present invention describes a DNA sequence which contains a base
CC sequence from bases 544-1758 in the sequence of (I) with 1992 base pairs,
CC or a DNA hybridisable with the DNA under stringent conditions and
CC encoding a polypeptide with hydroxylase activity on compactin at 6beta-
CC position. Also described: (1) DNA containing base sequences from bases
CC 544-1758 and from bases 1782-1970 in the sequence of (I) or a DNA
CC hybridisable with the DNA under stringent conditions and encoding a
CC polypeptide with hydroxylase activity on compactin at the 6beta-position;
CC (2) a polypeptide encoded by any of the DNA or containing an amino acid
CC sequence based on the polypeptide but with some amino acids deleted,
CC substituted or added and having hydroxylase activity on compactin at the
CC 6beta-position; (3) a recombinant DNA obtained by integrating with any of
CC the DNA; (4) a microorganism transferred with the recombinant DNA; (5) a
CC process for producing pravastatin by culturing the transformant
CC microorganism before isolating the culture liquor or cells, and addition
CC of compactin for reaction to give pravastatin for recovery; and (6)
CC Streptomyces sp. TM-6 (FERM BP-8002) or TM-7 (FERM BP-8003). (I) has
CC antilipaemic activity. The polypeptide and its encoded DNA are applicable
CC in constructing transformant microorganisms to synthesise pravastatin for
CC treating hyperlipidaemia. With the recombinant microorganisms,
CC pravastatin can be produced efficiently, with much less 6alpha
CC hydroxylated epimer formed. The present sequence represents a
CC Microtetraspora recticatena IFO14525 nucleotide sequence, which is given
CC in the exemplification of the present invention
XX
SQ Sequence 460 BP; 97 A; 119 C; 166 G; 78 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 8; Length 460;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 15 GCGTGCTTAACACATGCAAGTC 36

Search completed: April 7, 2006, 19:22:24
Job time : 224 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:15:09 ; Search time 1708.5 Seconds
(without alignments)
602.468 Million cell updates/sec

Title: US-10-697-802A-42
Perfect score: 22
Sequence: 1 gcgtgcttaacacatgcaagtc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	22	100.0	139	3	BP926260	BP926260 BP926260
c 2	22	100.0	273	7	CR476562	CR476562 CR476562
c 3	22	100.0	278	6	CD098227	CD098227 ME1-0019T
4	22	100.0	309	7	CR460297	CR460297 CR460297
c 5	22	100.0	313	1	AI903761	AI903761 IL-BT037-
c 6	22	100.0	356	6	CD122074	CD122074 ME1-0071G
7	22	100.0	456	3	BP924166	BP924166 BP924166
8	22	100.0	513	6	CA282280	CA282280 SCAGSD204
9	22	100.0	576	7	CN205661	CN205661 Tor6069 G
10	22	100.0	591	7	CN207299	CN207299 Tor7720 G
c 11	22	100.0	615	6	CD459102	CD459102 Fg08_05d1
12	22	100.0	641	6	CA285433	CA285433 SCEQSD107
13	22	100.0	650	7	CN204419	CN204419 Tor4810 G
14	22	100.0	657	7	CN208729	CN208729 Tor9226 G
c 15	22	100.0	663	6	CD096847	CD096847 ME1-0010T
c 16	22	100.0	722	6	CD164440	CD164440 ML1-0087T
17	22	100.0	725	7	CN204148	CN204148 Tor4539 G
18	22	100.0	740	6	CD164477	CD164477 ML1-0087T
19	22	100.0	744	6	CD164478	CD164478 ML1-0087T
c 20	22	100.0	887	10	CL693661	CL693661 PRI0162a_
c 21	21	95.5	654	9	BH578749	BH578749 BOGPY39TF
22	20.4	92.7	83	10	CW352030	CW352030 fsbb001f0
23	20.4	92.7	102	6	CA808630	CA808630 CA12LIO3I
24	20.4	92.7	120	7	CK054617	CK054617 63094rsic
25	20.4	92.7	128	9	BH567017	BH567017 BOHHZ63TR
26	20.4	92.7	149	10	CW422586	CW422586 fsbb001f1
27	20.4	92.7	151	1	AV637151	AV637151 AV637151
c 28	20.4	92.7	160	9	BH602265	BH602265 BOHAK51TF
29	20.4	92.7	162	8	DN477390	DN477390 altr303xd
c 30	20.4	92.7	164	9	CC728503	CC728503 OGWCG31TV
31	20.4	92.7	167	9	CC728495	CC728495 OGWCG31TH
32	20.4	92.7	172	6	CD831917	CD831917 BN40.061F
33	20.4	92.7	179	3	BP906725	BP906725 BP906725
34	20.4	92.7	179	3	BP907080	BP907080 BP907080
35	20.4	92.7	179	3	BP907569	BP907569 BP907569
36	20.4	92.7	179	3	BP907695	BP907695 BP907695
37	20.4	92.7	179	3	BP908232	BP908232 BP908232
38	20.4	92.7	180	9	BH718241	BH718241 BOMJU27TF
39	20.4	92.7	184	10	CW055265	CW055265 104_296_1
c 40	20.4	92.7	188	9	BZ500623	BZ500623 BONKS56TR
c 41	20.4	92.7	191	7	CK906470	CK906470 CSECS158E
42	20.4	92.7	195	8	CX944922	CX944922 DH0AGB7ZC
43	20.4	92.7	200	9	BZ483020	BZ483020 BONAM89TF
44	20.4	92.7	204	6	CA809833	CA809833 CA22LI01I
45	20.4	92.7	204	9	BH704086	BH704086 BOHVT76TF

ALIGNMENTS

RESULT 1
BP926260
LOCUS BP926260 139 bp mRNA linear EST 23-FEB-2005
DEFINITION BP926260 full-length enriched poplar cDNA library Populus nigra cDNA clone PnFL1-057_E19.f 5', mRNA sequence.
ACCESSION BP926260
VERSION BP926260.1 GI:60207890
KEYWORDS EST.
SOURCE Populus nigra
ORGANISM Populus nigra
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE 1 (bases 1 to 139)
AUTHORS Nanjo,T., Futamura,N., Nishiguchi,M., Igasaki,T., Shinozaki,K. and Shinohara,K.
TITLE Characterization of full-length enriched expressed sequence tags of stress-treated poplar leaves
JOURNAL Plant Cell Physiol. 45 (12), 1738-1748 (2004)
PUBMED 15653793
COMMENT Contact: Tokihiko Nanjo
Molecular and Cell Biology
Forestry and Forest Products Research Institute (FFPRI)
1 Matsunosato, Tsukuba, Ibaraki, 305-8687, Japan
Tel: 81-29-873-3211
Fax: 81-29-873-0507
Email: nanjo@affrc.go.jp.
FEATURES Location/Qualifiers
source 1. .139
/organism="Populus nigra"
/mol_type="mRNA"
/db_xref="taxon:3691"
/clone="PnFL1-057_E19.f"
/sex="female"
/tissue_type="leaf"
/dev_stage="juvenile"
/clone_lib="full-length enriched poplar cDNA library"
/note="synonym: Populus nigra var. italica"
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 139;
Best Local Similarity 100.0%; Pred. No. 2.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22

```
                    ||||||||||||||||||||
Db         44 GCGTGCTTAACACATGCAAGTC 65


RESULT 2
CR476562/c
LOCUS       CR476562                273 bp    mRNA    linear   EST 07-JUL-2004
DEFINITION  CR476562 Rat pBluescript Lion Rattus norvegicus cDNA clone
            LIONp463H07412 3', mRNA sequence.
ACCESSION   CR476562
VERSION     CR476562.1  GI:49902552
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 273)
  AUTHORS   Henrich,J., Hermanns,J., Kranz,H., Loebbert,R., Schlueter,T.,
            Schuette,D., Weindel,M., Heil,O., Ebert,L., Neubert,P., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Rat ArrayTAG cDNA
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; LIONp463H07412.
            RZPDLIB;
            Rat ArrayTAG cDNA
            http://www.rzpd.de/cgi-
            bin/products/showLib.pl.cgi/response?libNo=463 Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            RP: CAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .273
                     /organism="Rattus norvegicus"
                     /mol_type="mRNA"
                     /db_xref="taxon:10116"
                     /clone="LIONp463H07412"
                     /lab_host="DH10B"
                     /clone_lib="Rat pBluescript Lion"
ORIGIN

  Query Match.            100.0%;  Score 22;  DB 7;  Length 273;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        259 GCGTGCTTAACACATGCAAGTC 238


RESULT 3
CD098227/c
LOCUS       CD098227                278 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ME1-0019T-V084-H03-U.B ME1-0019 Schistosoma mansoni cDNA clone
            ME1-0019T-V084-H03.B, mRNA sequence.
ACCESSION   CD098227
VERSION     CD098227.1  GI:34648701
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 278)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
  PUBMED    12973350
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ME1-0019T-V084  row: 3  column: H.
FEATURES             Location/Qualifiers
     source          1. .278
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="ME1-0019T-V084-H03.B"
                     /sex="mixed pool"
                     /dev_stage="egg"
                     /lab_host="Mus musculus"
                     /clone_lib="ME1-0019"
                     /note="Vector: pGEM T-easy"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 278;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        235 GCGTGCTTAACACATGCAAGTC 214


RESULT 4
CR460297
LOCUS       CR460297                309 bp    mRNA    linear   EST 01-JUL-2004
DEFINITION  CR460297 Rat pBluescript Lion Rattus norvegicus cDNA clone
            LIONp463B04397 3', mRNA sequence.
ACCESSION   CR460297
VERSION     CR460297.1  GI:49592646
KEYWORDS    EST.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 309)
  AUTHORS   Henrich,J., Hermanns,J., Kranz,H., Loebbert,R., Schlueter,T.,
            Schuette,D., Weindel,M., Heil,O., Ebert,L., Neubert,P., Peters,M.,
            Radelof,U., Schneider,D. and Korn,B.
  TITLE     Rat ArrayTAG cDNA
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Email: www.rzpd.de
            RZPD; LIONp463B04397.
            RZPDLIB;
            Rat ArrayTAG cDNA
```

```
            http://www.rzpd.de/cgi-
            bin/products/showLib.pl.cgi/response?libNo=463 Contact: Inge Arlart
            RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
            Heubnerweg 6, D-14059 Berlin, Germany
            Tel: +49 30 32639 100
            Fax: +49 30 32639 111
            www.rzpd.de
            This clone is available royalty-free from RZPD;
            contact RZPD (clone@rzpd.de) for further information. Seq primer:
            RP: CAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .309
                      /organism="Rattus norvegicus"
                      /mol_type="mRNA"
                      /db_xref="taxon:10116"
                      /clone="LIONp463B04397"
                      /lab_host="DH10B"
                      /clone_lib="Rat pBluescript Lion"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 7;  Length 309;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         41 GCGTGCTTAACACATGCAAGTC 62


RESULT 5
AI903761/c
LOCUS       AI903761                313 bp    mRNA    linear   EST 30-MAR-2000
DEFINITION  IL-BT037-211198-005 BT037 Homo sapiens cDNA, mRNA sequence.
ACCESSION   AI903761
VERSION     AI903761.1  GI:6494148
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 313)
  AUTHORS   Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,
            Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,
            Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H.,
            Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,
            O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and
            Simpson,A.J.
  TITLE     Shotgun sequencing of the human transcriptome with ORF expressed
            sequence tags
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
   PUBMED   10737800
COMMENT     Contact: Simpson A.J.G.
            Laboratory of Cancer Genetics
            Ludwig Institute for Cancer Research
            Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
            Brazil
            Tel: +55-11-2704922
            Fax: +55-11-2707001
            Email: asimpson@ludwig.org.br
            This sequence was derived from the FAPESP/LICR Human Cancer Genome
            Project. This entry can be seen in the following URL
            (http://www.ludwig.org.br/seq/gethtml.pl?t1=IL&t2=IL-BT037-005.html
            &t3=211198&t4=1)
            Seq primer: puc 18 forward.
FEATURES              Location/Qualifiers
     source           1. .313
                      /organism="Homo sapiens"
                      /mol_type="mRNA"
                      /db_xref="taxon:9606"
                      /sex="female"
                      /dev_stage="Adult"
                      /clone_lib="BT037"
                      /note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:
                      SmaI; A mini-library was made by cloning products derived
                      from  ORESTES PCR (U.S. Letters Patent application No.
                      196,716 -  Ludwig Institute for Cancer Research) profiles
                      into  the pUC 18 vector. Reverse transcription of tissue
                      mRNA and cDNA amplification were performed under low
                      stringency conditions."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 313;
  Best Local Similarity   100.0%;  Pred. No. 2.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        270 GCGTGCTTAACACATGCAAGTC 249


RESULT 6
CD122074/c
LOCUS       CD122074                356 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ME1-0071G-A160-E04-1.B ME1-0071 Schistosoma mansoni cDNA clone
            ME1-0071G-A160-E04.B, mRNA sequence.
ACCESSION   CD122074
VERSION     CD122074.1  GI:34660126
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 356)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
   PUBMED   12973350
COMMENT     Other_ESTs: ME1-0071G-A160-E04-2.B
            Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ME1-0071G-A160  row: 4  column: E.
FEATURES              Location/Qualifiers
     source           1. .356
                      /organism="Schistosoma mansoni"
                      /mol_type="mRNA"
                      /db_xref="taxon:6183"
                      /clone="ME1-0071G-A160-E04.B"
                      /sex="mixed pool"
                      /dev_stage="egg"
                      /lab_host="Mus musculus"
                      /clone_lib="ME1-0071"
                      /note="Vector: pGEM T-easy"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 356;
  Best Local Similarity   100.0%;  Pred. No. 2.5;
```

```
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        310 GCGTGCTTAACACATGCAAGTC 289


RESULT 7
BP924166
LOCUS       BP924166                456 bp    mRNA    linear   EST 23-FEB-2005
DEFINITION  BP924166 full-length enriched poplar cDNA library Populus nigra
            cDNA clone PnFL1-029_B19.f 5', mRNA sequence.
ACCESSION   BP924166
VERSION     BP924166.1  GI:60205608
KEYWORDS    EST.
SOURCE      Populus nigra
  ORGANISM  Populus nigra
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eurosids I; Malpighiales; Salicaceae; Saliceae; Populus.
REFERENCE   1  (bases 1 to 456)
  AUTHORS   Nanjo,T., Futamura,N., Nishiguchi,M., Igasaki,T., Shinozaki,K. and
            Shinohara,K.
  TITLE     Characterization of full-length enriched expressed sequence tags of
            stress-treated poplar leaves
  JOURNAL   Plant Cell Physiol. 45 (12), 1738-1748 (2004)
   PUBMED   15653793
COMMENT     Contact: Tokihiko Nanjo
            Molecular and Cell Biology
            Forestry and Forest Products Research Institute (FFPRI)
            1 Matsunosato, Tsukuba, Ibaraki, 305-8687, Japan
            Tel: 81-29-873-3211
            Fax: 81-29-873-0507
            Email: nanjo@affrc.go.jp.
FEATURES             Location/Qualifiers
     source          1. .456
                     /organism="Populus nigra"
                     /mol_type="mRNA"
                     /db_xref="taxon:3691"
                     /clone="PnFL1-029_B19.f"
                     /sex="female"
                     /tissue_type="leaf"
                     /dev_stage="juvenile"
                     /clone_lib="full-length enriched poplar cDNA library"
                     /note="synonym: Populus nigra var. italica"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 456;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         43 GCGTGCTTAACACATGCAAGTC 64


RESULT 8
CA282280
LOCUS       CA282280                513 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  SCAGSD2042H09.g SD2 Saccharum officinarum cDNA clone SCAGSD2042H09
            5', mRNA sequence.
ACCESSION   CA282280
VERSION     CA282280.1  GI:36013534
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 513)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
            http://www.bcccenter.fcav.unesp.br
            Plate: 042  row: H  column: 09
            Seq primer: T7 Promoter Primer.
FEATURES             Location/Qualifiers
     source          1. .513
                     /organism="Saccharum officinarum"
                     /mol_type="mRNA"
                     /db_xref="taxon:4547"
                     /clone="SCAGSD2042H09"
                     /lab_host="DH10B"
                     /clone_lib="SD2"
                     /note="Organ: Developing seeds (small insert library);
                     Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                     unidirectional cDNA library generated from [Developing
                     seeds (small  insert library)]. cDNA was prepared from
                     polyA+ mRNA using SuperScript  Plasmid System Kit
                     (Invitrogen). The double-strand cDNAs were  fractionated
                     in a sepharose CL-2B 40cm-columns and fragments sizing
                     between 0.8 and 1.5 Kb were directionally cloned into the
                     vector. Details  of each source of RNA and library
                     construction can be obtained at
                     http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 6;  Length 513;
  Best Local Similarity   100.0%;  Pred. No. 2.6;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         12 GCGTGCTTAACACATGCAAGTC 33


RESULT 9
CN205661
LOCUS       CN205661                576 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  Tor6069 Gametophyte rehydration Library Tortula ruralis cDNA, mRNA
            sequence.
ACCESSION   CN205661
VERSION     CN205661.1  GI:46902392
KEYWORDS    EST.
SOURCE      Tortula ruralis
  ORGANISM  Tortula ruralis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Dicranidae; Pottiales; Pottiaceae; Tortula.
REFERENCE   1  (bases 1 to 576)
  AUTHORS   Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauget,S.A. and Payton,P.R.
  TITLE     The rehydration transcriptome of the desiccation-tolerant bryophyte
            Tortula ruralis: transcript classification and analysis
  JOURNAL   BMC Genomics 5 (1), 89 (2004)
   PUBMED   15546486
COMMENT     Contact: Oliver Melvin J
            Plant Stress Lab
            USDA-ARS
            3810 4th St, Lubbock, TX 79415, USA
            Tel: 806-749-5560
            Fax: 806-723-5272
            Email: moliver@lbk.ars.usda.gov
            PCR PRimers
            FORWARD: GTTTTCCCAGTCACGAC
            BACKWARD: CAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .576
```

```
                     /organism="Tortula ruralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:38588"
                     /clone_lib="Gametophyte rehydration Library"
                     /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
                     SalI; Site_2: NotI"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 7;  Length 576;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         85 GCGTGCTTAACACATGCAAGTC 106


RESULT 10
CN207299
LOCUS       CN207299                591 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  Tor7720 Gametophyte rehydration Library Tortula ruralis cDNA, mRNA
            sequence.
ACCESSION   CN207299
VERSION     CN207299.1  GI:46904030
KEYWORDS    EST.
SOURCE      Tortula ruralis
  ORGANISM  Tortula ruralis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Dicranidae; Pottiales; Pottiaceae; Tortula.
REFERENCE   1  (bases 1 to 591)
  AUTHORS   Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauget,S.A. and Payton,P.R.
  TITLE     The rehydration transcriptome of the desiccation-tolerant bryophyte
            Tortula ruralis: transcript classification and analysis
  JOURNAL   BMC Genomics 5 (1), 89 (2004)
   PUBMED   15546486
COMMENT     Contact: Oliver Melvin J
            Plant Stress Lab
            USDA-ARS
            3810 4th St, Lubbock, TX 79415, USA
            Tel: 806-749-5560
            Fax: 806-723-5272
            Email: moliver@lbk.ars.usda.gov
            PCR PRimers
            FORWARD: GTTTTCCCAGTCACGAC
            BACKWARD: CAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1..591
                     /organism="Tortula ruralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:38588"
                     /clone_lib="Gametophyte rehydration Library"
                     /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
                     SalI; Site_2: NotI"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 7;  Length 591;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        259 GCGTGCTTAACACATGCAAGTC 280


RESULT 11
CD459102/c
LOCUS       CD459102                615 bp    mRNA    linear   EST 14-JUN-2004
DEFINITION  Fg08_05d10_A Fg08_AAFC_ECORC_Fusarium_graminearum_complex_substrate
            Gibberella zeae cDNA clone Fg08_05d10, mRNA sequence.
ACCESSION   CD459102
VERSION     CD459102.3  GI:48688875
KEYWORDS    EST.
SOURCE      Gibberella zeae
  ORGANISM  Gibberella zeae
            Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
            Hypocreomycetidae; Hypocreales; Nectriaceae; Gibberella.
REFERENCE   1  (bases 1 to 615)
  AUTHORS   Watson,R.J., Heys,R., Chapados,J., Couroux,P., Harris,L.J.,
            Hattori,J., Lacroix,C., Ouellet,T., Robert,L.S., Singh,J.A.,
            Sprott,D. and Tinker,N.A.
  TITLE     A cDNA library prepared from Fusarium graminearum grown on a
            complex plant substrate
  JOURNAL   Unpublished (2003)
COMMENT     On Jun 3, 2003 this sequence version replaced gi:40466770.
            Contact: Watson, Robert.J.
            Eastern Cereal and Oilseed Research Centre
            Agriculture and Agri-food Canada
            Bldg. 20, Central Experimental Farm, Ottawa, Ontario, K1A 0C6,
            CANADA
            Tel: (613) 759-1655
            Fax: (613) 759-1701
            Email: watsonrj@agr.gc.ca.
FEATURES             Location/Qualifiers
     source          1. .615
                     /organism="Gibberella zeae"
                     /mol_type="mRNA"
                     /strain="DAOM 180378"
                     /db_xref="taxon:5518"
                     /clone="Fg08_05d10"
                     /tissue_type="Mycelium"
                     /dev_stage="Asexual"
                     /lab_host="E. coli DH10B"
                     /clone_lib="Fg08_AAFC_ECORC_Fusarium_graminearum_complex_s
                     ubstrate"
                     /note="Vector: pBluescript II+; Site_1: EcoRI; Site_2:
                     XhoI; Fusarium graminearum grown on a complex plant
                     substrate-- wheat leaves treated to remove most of the low
                     molecular weight, water-soluble components."
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 615;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        381 GCGTGCTTAACACATGCAAGTC 360


RESULT 12
CA285433
LOCUS       CA285433                641 bp    mRNA    linear   EST 26-SEP-2003
DEFINITION  SCEQSD1076H11.g SD1 Saccharum officinarum cDNA clone SCEQSD1076H11
            5', mRNA sequence.
ACCESSION   CA285433
VERSION     CA285433.1  GI:36025694
KEYWORDS    EST.
SOURCE      Saccharum officinarum
  ORGANISM  Saccharum officinarum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum
            complex.
REFERENCE   1  (bases 1 to 641)
  AUTHORS   Vettore,A.L., da Silva,F.R., Kemper,E.L. and Arruda,P.
  TITLE     The libraries that made SUCEST
  JOURNAL   Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
COMMENT     Contact: Arruda P
            Centro de Biologia Molecular e Engenharia Genetica
            Universidade Estadual de Campinas
            Caixa Postal 6010, 13083-970, Campinas SP, Brazil
            Tel: 55 19 3788 1137
            Fax: 55 19 3788 1089
            Email: parruda@unicamp.br
            Clone distribution: clone distribution information can be found
            through the Brazilian Clone Collection Center (BCCC) at
```

```
            http://www.bcccenter.fcav.unesp.br
            Plate: 076  row: H  column: 11
            Seq primer: T7 Promoter Primer.
FEATURES              Location/Qualifiers
     source           1. .641
                      /organism="Saccharum officinarum"
                      /mol_type="mRNA"
                      /db_xref="taxon:4547"
                      /clone="SCEQSD1076H11"
                      /lab_host="DH10B"
                      /clone_lib="SD1"
                      /note="Organ: Developing seeds (large insert library);
                      Vector: pSport1; Site_1: SalI; Site_2: NotI; An
                      unidirectional cDNA library generated from [Developing
                      seeds (large  insert library)]. cDNA was prepared from
                      polyA+ mRNA using SuperScript  Plasmid System Kit
                      (Invitrogen). The double-strand cDNAs were  fractionated
                      in a sepharose CL-2B 40cm-columns and fragments sizing
                      between 0.8 and 1.5 Kb were directionally cloned into the
                      vector. Details  of each source of RNA and library
                      construction can be obtained at
                      http://sucest.lad.ic.unicamp.br/public"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 6;  Length 641;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         12 GCGTGCTTAACACATGCAAGTC 33


RESULT 13
CN204419
LOCUS       CN204419                650 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  Tor4810 Gametophyte rehydration Library Tortula ruralis cDNA, mRNA
            sequence.
ACCESSION   CN204419
VERSION     CN204419.1  GI:46901150
KEYWORDS    EST.
SOURCE      Tortula ruralis
  ORGANISM  Tortula ruralis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Dicranidae; Pottiales; Pottiaceae; Tortula.
REFERENCE   1  (bases 1 to 650)
  AUTHORS   Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauget,S.A. and Payton,P.R.
  TITLE     The rehydration transcriptome of the desiccation-tolerant bryophyte
            Tortula ruralis: transcript classification and analysis
  JOURNAL   BMC Genomics 5 (1), 89 (2004)
   PUBMED   15546486
COMMENT     Contact: Oliver Melvin J
            Plant Stress Lab
            USDA-ARS
            3810 4th St, Lubbock, TX 79415, USA
            Tel: 806-749-5560
            Fax: 806-723-5272
            Email: moliver@lbk.ars.usda.gov
            PCR PRimers
            FORWARD: GTTTTCCCAGTCACGAC
            BACKWARD: CAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .650
                      /organism="Tortula ruralis"
                      /mol_type="mRNA"
                      /db_xref="taxon:38588"
                      /clone_lib="Gametophyte rehydration Library"
                      /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
                      SalI; Site_2: NotI"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 7;  Length 650;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db         18 GCGTGCTTAACACATGCAAGTC 39


RESULT 14
CN208729
LOCUS       CN208729                657 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  Tor9226 Gametophyte rehydration Library Tortula ruralis cDNA, mRNA
            sequence.
ACCESSION   CN208729
VERSION     CN208729.1  GI:46905460
KEYWORDS    EST.
SOURCE      Tortula ruralis
  ORGANISM  Tortula ruralis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Dicranidae; Pottiales; Pottiaceae; Tortula.
REFERENCE   1  (bases 1 to 657)
  AUTHORS   Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauget,S.A. and Payton,P.R.
  TITLE     The rehydration transcriptome of the desiccation-tolerant bryophyte
            Tortula ruralis: transcript classification and analysis
  JOURNAL   BMC Genomics 5 (1), 89 (2004)
   PUBMED   15546486
COMMENT     Contact: Oliver Melvin J
            Plant Stress Lab
            USDA-ARS
            3810 4th St, Lubbock, TX 79415, USA
            Tel: 806-749-5560
            Fax: 806-723-5272
            Email: moliver@lbk.ars.usda.gov
            PCR PRimers
            FORWARD: GTTTTCCCAGTCACGAC
            BACKWARD: CAGGAAACAGCTATGAC.
FEATURES              Location/Qualifiers
     source           1. .657
                      /organism="Tortula ruralis"
                      /mol_type="mRNA"
                      /db_xref="taxon:38588"
                      /clone_lib="Gametophyte rehydration Library"
                      /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
                      SalI; Site_2: NotI"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 7;  Length 657;
  Best Local Similarity  100.0%;  Pred. No. 2.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 GCGTGCTTAACACATGCAAGTC 22
              ||||||||||||||||||||||
Db        224 GCGTGCTTAACACATGCAAGTC 245


RESULT 15
CD096847/c
LOCUS       CD096847                663 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ME1-0010T-M117-G11-U.G ME1-0010 Schistosoma mansoni cDNA clone
            ME1-0010T-M117-G11.G, mRNA sequence.
ACCESSION   CD096847
VERSION     CD096847.1  GI:34647360
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 663)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
```

Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T., Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M., Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
TITLE Transcriptome analysis of the acoelomate human parasite Schistosoma mansoni
JOURNAL Nat. Genet. 35 (2), 148-157 (2003)
PUBMED 12973350
COMMENT Contact: Dr. Sergio Verjovski-Almeida
Departamento de Bioquimica
Instituto de Quimica - Universidade de Sao Paulo
Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP, Brasil
Tel: +55-11-3091-2173
Fax: +55-11-3091-2186
Email: verjo@iq.usp.br
This sequence was derived from the FAPESP Schistosoma mansoni EST Genome Project. All sequences in the project were assembled and annotated. This entry and all the assembled sequences can be seen in the following URL http://bioinfo.iq.usp.br/schisto/
Plate: ME1-0010T-M117 row: 11 column: G.

FEATURES Location/Qualifiers
source 1. .663
/organism="Schistosoma mansoni"
/mol_type="mRNA"
/db_xref="taxon:6183"
/clone="ME1-0010T-M117-G11.G"
/sex="mixed pool"
/dev_stage="egg"
/lab_host="Mus musculus"
/clone_lib="ME1-0010"
/note="Vector: pGEM T-easy"

ORIGIN

Query Match 100.0%; Score 22; DB 6; Length 663;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTGCTTAACACATGCAAGTC 22
||||||||||||||||||||||
Db 642 GCGTGCTTAACACATGCAAGTC 621

Search completed: April 7, 2006, 20:19:34
Job time : 1718.5 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:08:28 ; Search time 1183 Seconds
(without alignments)
1057.106 Million cell updates/sec

Title: US-10-697-802A-82
Perfect score: 22
Sequence: 1 tcctcctgatatctgcgcattc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_in:*
3: gb_env:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pr:*
9: gb_ro:*
10: gb_sts:*
11: gb_sy:*
12: gb_un:*
13: gb_vi:*
14: gb_htg:*
15: gb_pl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.		Score	Query Match %	Length	DB	ID	Description
c	1	22	100.0	163	3	UAU85181	U85181 Unidentifie
c	2	22	100.0	185	3	UAU85175	U85175 Unidentifie
c	3	22	100.0	189	3	UAU85191	U85191 Unidentifie
c	4	22	100.0	210	3	AF045838	AF045838 Unculture
c	5	22	100.0	213	3	AF045837	AF045837 Unculture
c	6	22	100.0	216	3	AF045840	AF045840 Unculture
c	7	22	100.0	236	3	UAU85185	U85185 Unidentifie
c	8	22	100.0	259	3	UAU85184	U85184 Unidentifie
	9	22	100.0	270	3	AY897639	AY897639 Unculture
c	10	22	100.0	291	1	AI16S3	X67455 A.israelii
c	11	22	100.0	298	1	SSP270383	AJ270383 Saccharom
c	12	22	100.0	298	1	SSP270384	AJ270384 Saccharom
c	13	22	100.0	298	3	UAU85186	U85186 Unidentifie
c	14	22	100.0	299	1	SSP270378	AJ270378 Saccharom
c	15	22	100.0	311	3	UAU85174	U85174 Unidentifie
c	16	22	100.0	316	3	AY886739	AY886739 Unculture
c	17	22	100.0	316	3	UAU85180	U85180 Unidentifie
c	18	22	100.0	317	1	SSP270373	AJ270373 Saccharom
c	19	22	100.0	318	1	SSP270374	AJ270374 Saccharom
c	20	22	100.0	319	3	UAU85179	U85179 Unidentifie
c	21	22	100.0	319	3	UAU85188	U85188 Unidentifie
c	22	22	100.0	320	3	AY886733	AY886733 Unculture
c	23	22	100.0	320	3	UAU85177	U85177 Unidentifie
c	24	22	100.0	326	3	AY897685	AY897685 Unculture
c	25	22	100.0	330	3	AF143761	AF143761 Unculture
c	26	22	100.0	331	1	MCC16SRNA	X86005 M.curtisii
c	27	22	100.0	331	1	MCH16SRNA	X86006 M.curtisii
c	28	22	100.0	331	1	MMRRNA16S	X86004 M.mulieris
c	29	22	100.0	333	3	UAU85178	U85178 Unidentifie
c	30	22	100.0	336	3	UAU85176	U85176 Unidentifie
c	31	22	100.0	340	1	AF250414	AF250414 Microbact
c	32	22	100.0	340	1	AF488639	AF488639 Gram-posi
c	33	22	100.0	340	3	BSPX91529	X91529 Uncultured
c	34	22	100.0	341	3	UAU85187	U85187 Unidentifie
	35	22	100.0	346	3	AF240478	AF240478 Unculture
c	36	22	100.0	346	3	UNC225376	AJ225376 Terrabact
c	37	22	100.0	348	3	BSPS13	Z69317 Uncultured
c	38	22	100.0	348	3	UNC225370	AJ225370 Terrabact
	39	22	100.0	351	3	AY886688	AY886688 Unculture
c	40	22	100.0	353	1	AY267529	AY267529 Microbact
c	41	22	100.0	353	3	UNC225379	AJ225379 Terrabact
c	42	22	100.0	355	1	AJ630198	AJ630198 Micrococc
c	43	22	100.0	356	3	UNC225341	AJ225341 Terrabact
c	44	22	100.0	358	1	SSP270379	AJ270379 Saccharom
c	45	22	100.0	358	1	AY827935	AY827935 Actinomyc

ALIGNMENTS

RESULT 1
UAU85181/c
LOCUS UAU85181 163 bp DNA linear ENV 03-MAY-2004
DEFINITION Unidentified actinomycetales clone ACK-M17 16S ribosomal RNA gene, partial sequence.
ACCESSION U85181
VERSION U85181.1 GI:2281365
KEYWORDS ENV.
SOURCE uncultured actinomycete
ORGANISM uncultured actinomycete
Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales; environmental samples.
REFERENCE 1 (bases 1 to 163)
AUTHORS Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
TITLE Bacterial diversity in Adirondack mountain lakes as revealed by 16S rRNA gene sequences
JOURNAL Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
PUBMED 9212443
REFERENCE 2 (bases 1 to 163)
AUTHORS Methe,B.A.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES Location/Qualifiers
source 1. .163
/organism="uncultured actinomycete"
/mol_type="genomic DNA"
/db_xref="taxon:100235"
/clone="ACK-M17"
/environmental_sample
/note="Uncultivated organism in integrated epilimnetic sample from Moss Lake, NY, USA"
rRNA <1. .>163
/product="16S ribosomal RNA"
ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 163;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22

```
                    ||||||||||||||||||||||
Db        109 TCCTCCTGATATCTGCGCATTC 88


RESULT 2
UAU85175/c
LOCUS       UAU85175                185 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-C53 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85175
VERSION     U85175.1  GI:2281359
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 185)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
   PUBMED   9212443
REFERENCE   2  (bases 1 to 185)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .185
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-C53"
                     /environmental_sample
                     /note="Uncultivated organism in integrated epilimnetic
                     sample from Carry Pond, NY, USA"
     rRNA            <1. .>185
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 185;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         63 TCCTCCTGATATCTGCGCATTC 42


RESULT 3
UAU85191/c
LOCUS       UAU85191                189 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-M2 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85191
VERSION     U85191.1  GI:2281375
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 189)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
   PUBMED   9212443
REFERENCE   2  (bases 1 to 189)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .189
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-M2"
                     /environmental_sample
                     /note="Uncultivated organism in integrated epilimnetic
                     sample from Moss Lake, NY, USA"
     rRNA            <1. .>189
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 189;
  Best Local Similarity   100.0%;  Pred. No. 3.8e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        109 TCCTCCTGATATCTGCGCATTC 88


RESULT 4
AF045838/c
LOCUS       AF045838                210 bp    DNA     linear   ENV 04-MAY-2004
DEFINITION  Uncultured bacterium clone L6a 16S ribosomal RNA gene, partial
            sequence.
ACCESSION   AF045838
VERSION     AF045838.1  GI:4105467
KEYWORDS    ENV.
SOURCE      uncultured bacterium
  ORGANISM  uncultured bacterium
            Bacteria; environmental samples.
REFERENCE   1  (bases 1 to 210)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Characterization of the bacterial population in the blood of the
            ascidian, Ascidia interrupta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 210)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-1998) Biological Sciences, UNCW, 601 South
            College Rd., Wilmington, NC 28403, USA
FEATURES             Location/Qualifiers
     source          1. .210
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /specific_host="Ascidia interrupta"
                     /db_xref="taxon:77133"
                     /clone="L6a"
                     /environmental_sample
                     /note="PCR-amplified from bacteria isolated from ascidian
                     blood using bacterial 16S rDNA-specific primers"
     rRNA            <1. .>210
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 210;
  Best Local Similarity   100.0%;  Pred. No. 3.6e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         60 TCCTCCTGATATCTGCGCATTC 39


RESULT 5
AF045837/c
LOCUS       AF045837                213 bp    DNA     linear   ENV 04-MAY-2004
DEFINITION  Uncultured bacterium clone L3a 16S ribosomal RNA gene, partial
            sequence.
ACCESSION   AF045837
VERSION     AF045837.1  GI:4105466
KEYWORDS    ENV.
```

```
SOURCE      uncultured bacterium
  ORGANISM  uncultured bacterium
            Bacteria; environmental samples.
REFERENCE   1  (bases 1 to 213)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Characterization of the bacterial population in the blood of the
            ascidian, Ascidia interrupta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 213)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-1998) Biological Sciences, UNCW, 601 South
            College Rd., Wilmington, NC 28403, USA
FEATURES             Location/Qualifiers
     source          1. .213
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /specific_host="Ascidia interrupta"
                     /db_xref="taxon:77133"
                     /clone="L3a"
                     /environmental_sample
                     /note="PCR-amplified from bacteria isolated from ascidian
                     blood using bacterial 16S rDNA-specific primers"
     rRNA            <1. .>213
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 213;
  Best Local Similarity  100.0%;  Pred. No. 3.6e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         61 TCCTCCTGATATCTGCGCATTC 40


RESULT 6
AF045840/c
LOCUS       AF045840                216 bp    DNA     linear   ENV 04-MAY-2004
DEFINITION  Uncultured bacterium clone L8b 16S ribosomal RNA gene, partial
            sequence.
ACCESSION   AF045840
VERSION     AF045840.1  GI:4105469
KEYWORDS    ENV.
SOURCE      uncultured bacterium
  ORGANISM  uncultured bacterium
            Bacteria; environmental samples.
REFERENCE   1  (bases 1 to 216)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Characterization of the bacterial population in the blood of the
            ascidian, Ascidia interrupta
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 216)
  AUTHORS   Williams,K.P., Sizemore,R.K. and Bartl,S.
  TITLE     Direct Submission
  JOURNAL   Submitted (04-FEB-1998) Biological Sciences, UNCW, 601 South
            College Rd., Wilmington, NC 28403, USA
FEATURES             Location/Qualifiers
     source          1. .216
                     /organism="uncultured bacterium"
                     /mol_type="genomic DNA"
                     /specific_host="Ascidia interrupta"
                     /db_xref="taxon:77133"
                     /clone="L8b"
                     /environmental_sample
                     /note="PCR-amplified from bacteria isolated from ascidian
                     blood using bacterial 16S rDNA-specific primers"
     rRNA            <1. .>216
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 216;
  Best Local Similarity  100.0%;  Pred. No. 3.6e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         58 TCCTCCTGATATCTGCGCATTC 37


RESULT 7
UAU85185/c
LOCUS       UAU85185                236 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-C3 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85185
VERSION     U85185.1  GI:2281369
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 236)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
  PUBMED    9212443
REFERENCE   2  (bases 1 to 236)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .236
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-C3"
                     /environmental_sample
                     /note="Uncultivated organism in integrated epilimnetic
                     sample from Carry Pond, NY, USA"
     rRNA            <1. .>236
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 236;
  Best Local Similarity  100.0%;  Pred. No. 3.5e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         95 TCCTCCTGATATCTGCGCATTC 74


RESULT 8
UAU85184/c
LOCUS       UAU85184                259 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-DH8 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85184
VERSION     U85184.1  GI:2281368
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 259)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
  PUBMED    9212443
REFERENCE   2  (bases 1 to 259)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
```

```
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .259
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-DH8"
                     /environmental_sample
                     /note="Uncultivated organism in hypolimnetic sample from
                     Dart's Lake, NY, USA"
     rRNA            <1. .>259
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 259;
  Best Local Similarity   100.0%;  Pred. No. 3.4e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        103 TCCTCCTGATATCTGCGCATTC 82


RESULT 9
AY897639
LOCUS       AY897639                 270 bp    DNA     linear   ENV 31-MAR-2005
DEFINITION  Uncultured organism clone MB042613 small subunit ribosomal RNA
            gene, partial sequence.
ACCESSION   AY897639
VERSION     AY897639.1  GI:59894900
KEYWORDS    ENV.
SOURCE      uncultured organism
  ORGANISM  uncultured organism
            unclassified; environmental samples.
REFERENCE   1  (bases 1 to 270)
  AUTHORS   Angenent,L.T., Kelley,S.T., Amand,A.S., Pace,N.R. and
            Hernandez,M.T.
  TITLE     Molecular identification of potential pathogens in water and air of
            a hospital therapy pool
  JOURNAL   Proc. Natl. Acad. Sci. U.S.A. 102 (13), 4860-4865 (2005)
   PUBMED   15769858
REFERENCE   2  (bases 1 to 270)
  AUTHORS   Angenent,L.T., Kelley,S.T., St. Amand,A.L., Pace,N.R. and
            Hernandez,M.T.
  TITLE     Direct Submission
  JOURNAL   Submitted (19-JAN-2005) Department of Chemical Engineering and
            Environmental Engineering Science Program, Washington University in
            St. Louis, Washington University Campus Box 1180, St. Louis, MO
            63130, USA
FEATURES             Location/Qualifiers
     source          1. .270
                     /organism="uncultured organism"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:155900"
                     /clone="MB042613"
                     /environmental_sample
     rRNA            complement(<1. .>270)
                     /product="small subunit ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 270;
  Best Local Similarity   100.0%;  Pred. No. 3.3e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         70 TCCTCCTGATATCTGCGCATTC 91


RESULT 10
AI16S3/c
LOCUS       AI16S3                   291 bp    DNA     linear   BCT 06-JUN-2003
DEFINITION  A.israelii serotype 2 16S rRNA (part 3 of 4).
ACCESSION   X67455
VERSION     X67455.1  GI:38846
KEYWORDS    16S ribosomal RNA.
SOURCE      Actinomyces israelii
  ORGANISM  Actinomyces israelii
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Actinomycineae; Actinomycetaceae; Actinomyces.
REFERENCE   1  (bases 1 to 291)
  AUTHORS   Stackebrandt,E. and Charfreitag,O.
  TITLE     Partial 16S rRNA primary structure of five Actinomyces species:
            phylogenetic implications and development of an Actinomyces
            israelii-specific oligonucleotide probe
  JOURNAL   J. Gen. Microbiol. 136 (Pt 1), 37-43 (1990)
   PUBMED   1693659
FEATURES             Location/Qualifiers
     source          1. .291
                     /organism="Actinomyces israelii"
                     /mol_type="genomic DNA"
                     /strain="DSM 43322"
                     /db_xref="taxon:1659"
     rRNA            1. .291
                     /product="16S ribosomal RNA"
                     /note="see also x67453,x67454,x67456"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 1;  Length 291;
  Best Local Similarity   100.0%;  Pred. No. 3.2e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         39 TCCTCCTGATATCTGCGCATTC 18


RESULT 11
SSP270383/c
LOCUS       SSP270383                298 bp    DNA     linear   BCT 13-DEC-2000
DEFINITION  Saccharomonospora sp. 42-190 partial 16S rRNA gene, isolate 42-190.
ACCESSION   AJ270383
VERSION     AJ270383.1  GI:11863699
KEYWORDS    16S ribosomal RNA; 16S rRNA gene.
SOURCE      Saccharomonospora sp. 42-190
  ORGANISM  Saccharomonospora sp. 42-190
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Pseudonocardineae; Pseudonocardiaceae; Saccharomonospora.
REFERENCE   1
  AUTHORS   Salazar,O., Moron,R. and Genilloud,O.
  TITLE     New genus-specific primers for the PCR identification of members of
            the genus Saccharomonospora and evaluation of the microbial
            diversity of wild-type isolates of Saccharomonospora detected from
            soil DNAs
  JOURNAL   Int. J. Syst. Evol. Microbiol. 50 Pt 6, 2043-2055 (2000)
   PUBMED   11155979
REFERENCE   2  (bases 1 to 298)
  AUTHORS   Genilloud,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-1999) Genilloud O., Centro de Investigacion
            Basica, NPDD-Merck Research Labs., Merck, Sharp & Dohme de Espana,
            S.A., Josefa Valcarcel 34, Madrid, SPAIN
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Saccharomonospora sp. 42-190"
                     /mol_type="genomic DNA"
                     /isolate="42-190"
                     /db_xref="taxon:105466"
                     /country="Mexico"
     gene            1. .298
                     /gene="16S rRNA"
     rRNA            <1. .>298
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"
ORIGIN
```

```
  Query Match           100.0%;  Score 22;  DB 1;  Length 298;
  Best Local Similarity  100.0%;  Pred. No. 3.2e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        145 TCCTCCTGATATCTGCGCATTC 124


RESULT 12
SSP270384/c
LOCUS       SSP270384                298 bp    DNA     linear   BCT 13-DEC-2000
DEFINITION  Saccharomonospora sp. 42-193 partial 16S rRNA gene, isolate 42-193.
ACCESSION   AJ270384
VERSION     AJ270384.1  GI:11863700
KEYWORDS    16S ribosomal RNA; 16S rRNA gene.
SOURCE      Saccharomonospora sp. 42-193
  ORGANISM  Saccharomonospora sp. 42-193
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Pseudonocardineae; Pseudonocardiaceae; Saccharomonospora.
REFERENCE   1
  AUTHORS   Salazar,O., Moron,R. and Genilloud,O.
  TITLE     New genus-specific primers for the PCR identification of members of
            the genus Saccharomonospora and evaluation of the microbial
            diversity of wild-type isolates of Saccharomonospora detected from
            soil DNAs
  JOURNAL   Int. J. Syst. Evol. Microbiol. 50 Pt 6, 2043-2055 (2000)
   PUBMED   11155979
REFERENCE   2  (bases 1 to 298)
  AUTHORS   Genilloud,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-1999) Genilloud O., Centro de Investigacion
            Basica, NPDD-Merck Research Labs., Merck, Sharp & Dohme de Espana,
            S.A., Josefa Valcarcel 34, Madrid, SPAIN
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="Saccharomonospora sp. 42-193"
                     /mol_type="genomic DNA"
                     /isolate="42-193"
                     /db_xref="taxon:105467"
                     /country="Mexico"
     gene            1. .298
                     /gene="16S rRNA"
     rRNA            <1. .>298
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 1;  Length 298;
  Best Local Similarity  100.0%;  Pred. No. 3.2e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        145 TCCTCCTGATATCTGCGCATTC 124


RESULT 13
UAU85186/c
LOCUS       UAU85186                 298 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-C68 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85186
VERSION     U85186.1  GI:2281370
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 298)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
   PUBMED   9212443
REFERENCE   2  (bases 1 to 298)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .298
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-C68"
                     /environmental_sample
                     /note="Uncultivated organism in integrated epilimnetic
                     sample from Carry Pond, NY, USA"
     rRNA            <1. .>298
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 3;  Length 298;
  Best Local Similarity  100.0%;  Pred. No. 3.2e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        109 TCCTCCTGATATCTGCGCATTC 88


RESULT 14
SSP270378/c
LOCUS       SSP270378                299 bp    DNA     linear   BCT 13-DEC-2000
DEFINITION  Saccharomonospora sp. 42-161 partial 16S rRNA gene, isolate 42-161.
ACCESSION   AJ270378
VERSION     AJ270378.1  GI:11863694
KEYWORDS    16S ribosomal RNA; 16S rRNA gene.
SOURCE      Saccharomonospora sp. 42-161
  ORGANISM  Saccharomonospora sp. 42-161
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            Pseudonocardineae; Pseudonocardiaceae; Saccharomonospora.
REFERENCE   1
  AUTHORS   Salazar,O., Moron,R. and Genilloud,O.
  TITLE     New genus-specific primers for the PCR identification of members of
            the genus Saccharomonospora and evaluation of the microbial
            diversity of wild-type isolates of Saccharomonospora detected from
            soil DNAs
  JOURNAL   Int. J. Syst. Evol. Microbiol. 50 Pt 6, 2043-2055 (2000)
   PUBMED   11155979
REFERENCE   2  (bases 1 to 299)
  AUTHORS   Genilloud,O.
  TITLE     Direct Submission
  JOURNAL   Submitted (22-SEP-1999) Genilloud O., Centro de Investigacion
            Basica, NPDD-Merck Research Labs., Merck, Sharp & Dohme de Espana,
            S.A., Josefa Valcarcel 34, Madrid, SPAIN
FEATURES             Location/Qualifiers
     source          1. .299
                     /organism="Saccharomonospora sp. 42-161"
                     /mol_type="genomic DNA"
                     /isolate="42-161"
                     /db_xref="taxon:105461"
                     /country="Sri Lanka"
     gene            1. .299
                     /gene="16S rRNA"
     rRNA            <1. .>299
                     /gene="16S rRNA"
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match           100.0%;  Score 22;  DB 1;  Length 299;
  Best Local Similarity  100.0%;  Pred. No. 3.2e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        129 TCCTCCTGATATCTGCGCATTC 108


RESULT 15
UAU85174/c
LOCUS       UAU85174                311 bp    DNA     linear   ENV 03-MAY-2004
DEFINITION  Unidentified actinomycetales clone ACK-C67 16S ribosomal RNA gene,
            partial sequence.
ACCESSION   U85174
VERSION     U85174.1  GI:2281358
KEYWORDS    ENV.
SOURCE      uncultured actinomycete
  ORGANISM  uncultured actinomycete
            Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
            environmental samples.
REFERENCE   1  (bases 1 to 311)
  AUTHORS   Hiorns,W.D., Methe,B.A., Nierzwicki-Bauer,S.A. and Zehr,J.P.
  TITLE     Bacterial diversity in Adirondack mountain lakes as revealed by 16S
            rRNA gene sequences
  JOURNAL   Appl. Environ. Microbiol. 63 (7), 2957-2960 (1997)
   PUBMED   9212443
REFERENCE   2  (bases 1 to 311)
  AUTHORS   Methe,B.A.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-JAN-1997) Biology Department, Rensselaer Polytechnic
            Institute, 110 8th Street, Troy, NY 12180-3590, USA
FEATURES             Location/Qualifiers
     source          1. .311
                     /organism="uncultured actinomycete"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:100235"
                     /clone="ACK-C67"
                     /environmental_sample
                     /note="Uncultivated organism in integrated epilimnetic
                     sample from Carry Pond, NY, USA"
     rRNA            <1. .>311
                     /product="16S ribosomal RNA"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 311;
  Best Local Similarity   100.0%;  Pred. No. 3.1e+03;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        109 TCCTCCTGATATCTGCGCATTC 88


Search completed: April  7, 2006, 20:42:20
Job time : 1186 secs
```

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:01:48 ; Search time 220 Seconds
(without alignments)
666.469 Million cell updates/sec

Title: US-10-697-802A-82
Perfect score: 22
Sequence: 1 tcctcctgatatctgcgcattc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_21:*
1: geneseqn1980s:*
2: geneseqn1990s:*
3: geneseqn2000s:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2002bs:*
8: geneseqn2003as:*
9: geneseqn2003bs:*
10: geneseqn2003cs:*
11: geneseqn2003ds:*
12: geneseqn2004as:*
13: geneseqn2004bs:*
14: geneseqn2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	22	14	AEA22481	Aea22481 Acid-fast
c	2	22	100.0	567	13	ADR45486	Adr45486 16S rRNA
c	3	22	100.0	619	8	ABZ79780	Abz79780 Cellulomo
c	4	22	100.0	619	10	ADF65480	Adf65480 Novel alp
c	5	22	100.0	787	2	AAV43262	Aav43262 Partial 1
c	6	22	100.0	815	4	AAF89996	Aaf89996 Nucleotid
	7	22	100.0	1062	4	AAS59697	Aas59697 Propionib
	8	22	100.0	1062	8	ACF64626	Acf64626 Propionib
c	9	22	100.0	1135	14	ADZ67281	Adz67281 Frigoriba
c	10	22	100.0	1271	2	AAV24293	Aav24293 Mycobacte
c	11	22	100.0	1321	14	AEA22410	Aea22410 Mycobacte
c	12	22	100.0	1383	14	AEA22400	Aea22400 Mycobacte
c	13	22	100.0	1391	2	AAT45276	Aat45276 Corynebac
c	14	22	100.0	1415	14	AEA22413	Aea22413 Mycobacte
c	15	22	100.0	1416	14	AEA22416	Aea22416 Mycobacte
c	16	22	100.0	1421	14	AEA22411	Aea22411 Mycobacte
c	17	22	100.0	1421	14	AEA22402	Aea22402 Mycobacte
c	18	22	100.0	1431	12	ADK66476	Adk66476 Corynebac
c	19	22	100.0	1431	12	ADK66445	Adk66445 Corynebac
c	20	22	100.0	1439	14	AEA22403	Aea22403 Mycobacte
c	21	22	100.0	1449	2	AAQ37639	Aaq37639 Mycobacte
	22	22	100.0	1449	10	ADG44144	Adg44144 Unknown o
	23	22	100.0	1449	10	ADG17999	Adg17999 Unknown b
	24	22	100.0	1449	11	ADL27934	Adl27934 RA3 16S r
	25	22	100.0	1449	12	ADF47790	Adf47790 Unknown b
c	26	22	100.0	1449	14	AEA22405	Aea22405 Mycobacte
c	27	22	100.0	1452	13	ADR90573	Adr90573 M intrace
c	28	22	100.0	1452	14	AEA22408	Aea22408 Mycobacte
c	29	22	100.0	1454	14	AEA22401	Aea22401 Mycobacte
c	30	22	100.0	1455	14	AEA22412	Aea22412 Mycobacte
c	31	22	100.0	1456	14	ADZ67282	Adz67282 Frigoriba
c	32	22	100.0	1461	14	AEA22406	Aea22406 Mycobacte
c	33	22	100.0	1462	14	AEA22415	Aea22415 Mycobacte
c	34	22	100.0	1463	14	AEA22409	Aea22409 Mycobacte
c	35	22	100.0	1464	3	AAZ35571	Aaz35571 Mycobacte
c	36	22	100.0	1464	5	AAS11027	Aas11027 Mycobacte
c	37	22	100.0	1465	10	ADB61680	Adb61680 16S rRNA
c	38	22	100.0	1469	13	ADR90574	Adr90574 M kansasi
c	39	22	100.0	1472	13	ADR90572	Adr90572 M avium 1
c	40	22	100.0	1482	14	AEA22404	Aea22404 Mycobacte
c	41	22	100.0	1484	14	AEA22414	Aea22414 Mycobacte
c	42	22	100.0	1517	11	AEB80305	Aeb80305 Organic w
c	43	22	100.0	1524	4	AAS30719	Aas30719 Mycobacte
c	44	22	100.0	1527	14	AEA22407	Aea22407 Mycobacte
c	45	22	100.0	1536	10	ADB61681	Adb61681 16S rRNA

ALIGNMENTS

RESULT 1
AEA22481
ID AEA22481 standard; DNA; 22 BP.
XX
AC AEA22481;
XX
DT 25-AUG-2005 (first entry)
XX
DE Acid-fast bacterium reverse (AFB-r) 16S rDNA PCR primer SEQ ID NO:82.
XX
KW microorganism identification; 16S rDNA; 16S ribosomal DNA; PCR; primer;
KW ss.
XX
OS Synthetic.
XX
PN US2005130168-A1.
XX
PD 16-JUN-2005.
XX
PF 31-OCT-2003; 2003US-00697802.
XX
PR 31-OCT-2003; 2003US-00697802.
XX
PA (HANX/) HAN X.
PA (PHAM/) PHAM A S.
XX
PI Han X, Pham AS;
XX
DR WPI; 2005-424597/43.
XX
PT Determining a bacterium species comprises providing oligonucleotide
PT primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS Claim 2; SEQ ID NO 82; 74pp; English.
XX
CC The invention relates to a method (M1) for determining a bacterium
CC species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC extracting a genomic nucleotide from the bacterium to provide a
CC nucleotide template; (c) annealing a region of a nucleotide template to a
CC specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC complimentary fashion, the primer set designed to provide a product
CC having a predetermined size dictated by a complimentary primer set; (d)

CC amplifying the region of the nucleotide template to produce the product;
CC and (e) determining a species of a bacterium in a nucleotide sequence of
CC the product. Also described is an alternative method (M2) for determining
CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or
CC variations, or (iii) a first forward primer having consecutive bases of
CC an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC second reverse primer having consecutive bases of an UB-r of AEA22517-
CC AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC comparing the product from the specimen with a nucleotide sequence from a
CC database to determine the bacterium species present in the specimen. The
CC methods are useful for determining a bacterium species. The present
CC sequence represents a reverse PCR primer for amplifying 16S rDNA regions
CC of acid-fast bacterium (AFB), which is used in the exemplification of the
CC present invention.
XX
SQ Sequence 22 BP; 3 A; 8 C; 3 G; 8 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 22;
Best Local Similarity 100.0%; Pred. No. 0.74;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 1 TCCTCCTGATATCTGCGCATTC 22

RESULT 2
ADR45486/c
ID ADR45486 standard; DNA; 567 BP.
XX
AC ADR45486;
XX
DT 18-NOV-2004 (first entry)
XX
DE 16S rRNA gene 357f-518r region DNA fragment SeqID75.
XX
KW 357f-518r; 16S rRNA; beta proteobacterium; ammonia oxidising bacteria;
KW activated sludge; ammonia liquid treatment plant; chemical oxygen demand;
KW COD; reduction; nitrification; denitrifying; ds.
XX
OS Unidentified.
XX
PN JP2004242578-A.
XX
PD 02-SEP-2004.
XX
PF 13-FEB-2003; 2003JP-00035713.
XX
PR 13-FEB-2003; 2003JP-00035713.
XX
PA (YAWA) NIPPON STEEL CORP.
XX
DR WPI; 2004-620179/60.
XX
PT Novel DNA fragment of microorganisms existing in activated sludge of
PT ammonia liquid treatment plant, useful as index microorganisms for
PT evaluating nitrification or denitrifying capability of ammonia liquid.
XX
PS Claim 43; SEQ ID NO 75; 133pp; Japanese.
XX
CC This invention relates to a novel DNA fragment comprising the 357f-518r
CC region of the 16S rRNA gene of beta proteobacteria, belonging to the
CC ammonia oxidising bacteria group, or CFB Bacteroides where bacteria
CC exists in activated sludge of an ammonia liquid treatment plant and used
CC for chemical oxygen demand (COD) reduction. The invention is useful in
CC the identification of microorganisms as nitrification or denitrifying
CC index microorganisms for evaluating the nitrification or denitrifying
CC capability of ammonia liquid of the activated sludge by fluorescence in
CC situ hybridisation (FISH). The invention is also useful for developing
CC apparatus for the processing of ammonia liquid. The DNA fragment enables
CC evaluation of the nitrification or denitrifying capability of
CC microorganisms. The present sequence is that of a 16S rRNA gene 357f-518r
CC region of the invention.
XX
SQ Sequence 567 BP; 129 A; 127 C; 198 G; 112 T; 0 U; 1 Other;

Query Match 100.0%; Score 22; DB 13; Length 567;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 356 TCCTCCTGATATCTGCGCATTC 335

RESULT 3
ABZ79780/c
ID ABZ79780 standard; DNA; 619 BP.
XX
AC ABZ79780;
XX
DT 12-MAY-2003 (first entry)
XX
DE Cellulomonas sp. nucleotide sequence SEQ ID NO:8.
XX
KW Glycoprotein; Saccharomyces cerevisiae; yeast; acidic sugar-chain;
KW mannose-6-phosphate; lysosomal disease; nephrotropic; haemostatic;
KW lysozyme; human lysosomal enzyme deficiency; Fabry disease;
KW Gaucher's disease; lysosomal enzyme; gene; ds.
XX
OS Cellulomonas sp.
XX
PN WO2002103027-A1.
XX
PD 27-DEC-2002.
XX
PF 14-JUN-2002; 2002WO-JP005965.
XX
PR 14-JUN-2001; 2001JP-00180907.
XX
PA (NAAD-) NAT INST ADVANCED IND SCI & TECHNOLOGY.
PA (TOKM-) TOKYO METROPOLITAN ORG MEDICAL RES.
PA (TAKE/) TAKEUCHI Y.
XX
PI Takeuchi M, Chiba Y, Jigami Y, Sakuraba H, Kobayashi K;
XX
DR WPI; 2003-210100/20.
XX
PT Production of glycoproteins by culturing cells transformed with lysosomal
PT enzyme yeast sugar-chain synthase variant, applicable as labeling marker
PT for transporting lysozyme of cells and in drug compositions.
XX
PS Example 3; Page 59; 61pp; Japanese.
XX
CC The present invention describes a method (M1) for producing an active
CC glycoprotein with an acidic sugar-chain containing a mannose-6-phosphate
CC at its non-reducing terminal comprises using a yeast. Also described: (1)
CC the glycoproteins produced by (M1), having an acidic sugar-chain
CC containing mannose-6-phosphate at its non-reducing terminal; (2) drug
CC compositions for treating and/or preventing lysosomal diseases containing
CC the glycoproteins; and (3) producing active glycoproteins having a high-
CC mannose-type sugar-chain that contains a mannose-6-phosphate at its non-
CC reducing terminal by using yeast. The glycoprotein has nephrotropic and
CC haemostatic activities. The produced glycoprotein can be used as a
CC labeling marker for transporting lysozyme and in drug compositions to
CC treat human lysosomal enzyme deficiency e.g. Fabry disease and Gaucher's

```
CC   disease. The lysosomal enzyme can be produced in large quantities for use
CC   as efficacious drugs. The present sequence represents a Cellulomonas sp.
CC   nucleotide sequence, which is used in an example from the present
CC   invention. N.B. The present sequence is designated SEQ ID NO:7 on page 29
CC   but is given as SEQ ID NO:8 in the Sequence Listing
XX
SQ   Sequence 619 BP; 153 A; 146 C; 208 G; 112 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 8;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        286 TCCTCCTGATATCTGCGCATTC 265


RESULT 4
ADF65480/c
ID   ADF65480 standard; DNA; 619 BP.
XX
AC   ADF65480;
XX
DT   12-FEB-2004  (first entry)
XX
DE   Novel alpha-mannosidase related Cellulomonas DNA sequence SeqID3.
XX
KW   alpha-mannosidase; enzymological; hydrolysis;
KW   glycoprotein saccharide chain; mannose preparation; ds.
XX
OS   Cellulomonas sp.
XX
PN   JP2002369679-A.
XX
PD   24-DEC-2002.
XX
PF   14-JUN-2001; 2001JP-00180906.
XX
PR   14-JUN-2001; 2001JP-00180906.
XX
PA   (DOKU-) DOKURITSU GYOSEI HOJIN SANGYO GIJUTSU SO.
PA   (KIRI ) KIRIN BREWERY KK.
XX
DR   WPI; 2003-600993/57.
XX
PT   Alpha-mannosidase derived from Cellulomonas sp. SO-5 (FERM BP-7628) with
PT   potent enzymic activity on glycoprotein saccharide chain.
XX
PS   Example 3; SEQ ID NO 3; 16pp; Japanese.
XX
CC   This invention relates to a novel alpha-mannosidase which possesses
CC   specific enzymological properties. The enzyme has potent enzymatic
CC   activity (hydrolysis) on glycoprotein saccharide chains which may be
CC   useful in the preparation of mannose.
XX
SQ   Sequence 619 BP; 153 A; 146 C; 208 G; 112 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 10;  Length 619;
  Best Local Similarity   100.0%;  Pred. No. 1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        286 TCCTCCTGATATCTGCGCATTC 265


RESULT 5
AAV43262/c
ID   AAV43262 standard; DNA; 787 BP.
XX
AC   AAV43262;
XX
DT   26-OCT-1998  (first entry)
XX
DE   Partial 16s DNA sequence of Arthrobacter.
XX
KW   16s DNA sequence; vaccine; protection; farmed; salmonoid fish;
KW   Renibacterium salmoninarum; bacterial kidney disease; ss.
XX
OS   Arthrobacter sp.
XX
PN   WO9833884-A1.
XX
PD   06-AUG-1998.
XX
PF   28-JAN-1998;   98WO-GB000256.
XX
PR   30-JAN-1997;   97GB-00001897.
XX
PA   (AQUA-) AQUA HEALTH EURO LTD.
XX
PI   Griffiths SG,  Salonius K;
XX
DR   WPI; 1998-437441/37.
XX
PT   Immune stimulating agent or vaccine containing non-virulent Arthrobacter
PT   - useful for, e.g. protecting salmonoid fish against Renibacterium
PT   salmoninarum.
XX
PS   Claim 3; Page 11; 16pp; English.
XX
CC   The present sequence represents a partial 16s DNA sequence of
CC   Arthrobacter (ATTCC 55921). This strain of arthrobacter is used to
CC   produce the immune stimulating agent or vaccine of the inventrion.
CC   Arthrobacter (which shares surface antigens with R. salmoninarum)
CC   stimulates powerful specific and non-specific immunity, and since it can
CC   survive in macrophages ensures prolonged stimulation and protection. The
CC   products are used to protect farmed salmonoid fish against Renibacterium
CC   salmoninarum, the causative agent of bacterial kidney disease
XX
SQ   Sequence 787 BP; 179 A; 173 C; 268 G; 167 T; 0 U; 0 Other;

  Query Match             100.0%;  Score 22;  DB 2;  Length 787;
  Best Local Similarity   100.0%;  Pred. No. 1.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        696 TCCTCCTGATATCTGCGCATTC 675


RESULT 6
AAF89996/c
ID   AAF89996 standard; DNA; 815 BP.
XX
AC   AAF89996;
XX
DT   06-AUG-2001  (first entry)
XX
DE   Nucleotide sequence of a 16S rDNA sequence from an unknown organism.
XX
KW   Metabolic pathway operon; polyketide; polyketide antibiotic; 16 rDNA; ss.
XX
OS   Unidentified.
XX
PN   WO200140497-A2.
XX
PD   07-JUN-2001.
XX
PF   27-NOV-2000; 2000WO-FR003311.
XX
PR   29-NOV-1999;   99FR-00015032.
PR   07-JUN-2000; 2000US-0209800P.
XX
PA   (AVET ) AVENTIS PHARMA SA.
XX
```

PI Jeannin P, Pernodet J, Guerineau M, Simonet P, Courtois S;
PI Cappellano C, Francou F, Raynal A, Ball M, Sezonov G, Tuphile K;
PI Frostegard A;
XX
DR WPI; 2001-374849/39.
XX
PT Collection of nucleic acids from environmental samples, useful for
PT identifying e.g. genes encoding polyketide synthases and derived
PT antibiotics.
XX
PS Claim 76; Page 253-254; 356pp; French.
XX
CC The specification describes a method for the preparation of a collection
CC of nucleic acids from organisms in a soil sample. The method comprises
CC milling a dried sample to produce microparticles; suspending these in
CC liquid buffer; extraction of nucleic acids from the microparticle;
CC passing nucleic acid-containing solution through a molecular sieve;
CC passing nucleic acid-enriched fractions through an anion exchange
CC chromatography material; and recovering fractions containing purified
CC nucleic acids. The nucleic acids are sources for sequences that encode
CC either operons involved in a metabolic pathway (specifically polyketide
CC synthesis) or polypeptides, particularly for production of therapeutic or
CC agricultural compounds, especially polyketide antibiotics. AAF89979-
CC AAF90025 represent 16S rDNA sequences, which were isolated using the
CC method of the invention
XX
SQ Sequence 815 BP; 193 A; 194 C; 267 G; 161 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 4; Length 815;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 171 TCCTCCTGATATCTGCGCATTC 150

RESULT 7
AAS59697
ID AAS59697 standard; DNA; 1062 BP.
XX
AC AAS59697;
XX
DT 13-FEB-2002 (first entry)
XX
DE Propionibacterium acnes immunogenic protein encoding DNA #192.
XX
KW SAPHO syndrome; synovitis; acne; pustulosis; hypertosis; osteomyelitis;
KW uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;
KW inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;
KW dermatological; osteopathic; neuroprotectant; ds.
XX
OS Propionibacterium acnes.
XX
PN WO200181581-A2.
XX
PD 01-NOV-2001.
XX
PF 20-APR-2001; 2001WO-US012865.
XX
PR 21-APR-2000; 2000US-0199047P.
PR 02-JUN-2000; 2000US-0208841P.
PR 07-JUL-2000; 2000US-0216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skeiky YAW, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris.
XX
PS Claim 1; SEQ ID NO 192; 1069pp; English.
XX
CC Sequences AAS59506-AAS59804 represent DNA molecules encoding
CC Propionibacterium acnes immunogenic polypeptides. The proteins and their
CC associated DNA sequences are used in the treatment, prevention and
CC diagnosis of medical conditions caused by P. acnes. The disorders include
CC SAPHO syndrome (synovitis, acne, pustulosis, hypertosis and
CC osteomyelitis), uveitis and endophthalmitis. P. acnes is also involved in
CC infections of bone, joints and the central nervous system, however it is
CC particularly involved in the inflammatory lesions associated with acne
CC vulgaris. A method for detecting the presence or absence of P. acnes in a
CC patient comprises contacting a sample with a binding agent that binds to
CC the proteins of the invention and determining the amount of bound protein
CC in the sample. The polypeptides may be used as antigens in the production
CC of antibodies specific for P. acnes proteins. These antibodies can be
CC used to downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA). This sequence encodes the
CC polypeptides shown in AAU65867-AAU65877 and AAU67824-AAU67826. Note: The
CC sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 1062 BP; 212 A; 338 C; 304 G; 205 T; 0 U; 3 Other;

Query Match 100.0%; Score 22; DB 4; Length 1062;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 46 TCCTCCTGATATCTGCGCATTC 67

RESULT 8
ACF64626
ID ACF64626 standard; DNA; 1062 BP.
XX
AC ACF64626;
XX
DT 17-OCT-2003 (first entry)
XX
DE Propionibacterium acnes DNA contig sequence #192.
XX
KW Acne vulgaris; antiseborrhoeic; dermatological; antibacterial;
KW immunostimulant; immune response; vaccine; ds.
XX
OS Propionibacterium acnes.
XX
PN WO2003033515-A1.
XX
PD 24-APR-2003.
XX
PF 11-OCT-2002; 2002WO-US032727.
XX
PR 15-OCT-2001; 2001US-00978825.
XX
PA (CORI-) CORIXA CORP.
XX
PI Mitcham JL, Skeiky YAW, Persing DH, Bhatia A, Maisonneuve JL;
PI Zhang Y, Wang S, Jen S, Lodes MJ, Benson DR, Jones R, Carter D;
PI Barth B, Vallieve-Douglass J;
XX
DR WPI; 2003-381789/36.
XX
PT New Propionibacterium acnes polypeptides and polynucleotides encoding the
PT polypeptide, useful for diagnosing, preventing or treating acne vulgaris,
PT or for stimulating an immune response specific for a P. acnes protein.
XX
PS Claim 1; SEQ ID NO 192; 1481pp; English.

```
XX
CC   The invention relates to an isolated polynucleotide (ACF64435-ACF64733)
CC   encoding a Propionibacterium acnes protein. The invention also relates to
CC   polypeptides encoded by the polynucleotides (ABM35624-ABM64536) and to
CC   immunogenic fragments of P. acnes polypeptides. The invention
CC   additionally encompasses expression vectors and host cells comprising a
CC   polynucleotide of the invention; antibodies against polypeptides of the
CC   invention; fusion proteins comprising a polypeptide of the invention; a
CC   method for stimulating an immune response specific for a P. acnes
CC   polypeptide and an isolated T cell population comprising T cells prepared
CC   via this method; a vaccine composition (comprising P. acnes polypeptides,
CC   polynucleotides, antibodies, fusion proteins, T cell populations, or
CC   antigen-presenting cells that express the polypeptide); a method and kit
CC   for detecting or determining the presence or absence of P. acnes in a
CC   patient; and a method for inhibiting the development of P. acnes in a
CC   patient. The P. acnes polypeptides, polynucleotides, antibodies, fusion
CC   proteins, T cell populations or antigen-presenting cells that express the
CC   polypeptides are useful for diagnosing, preventing or treating acne
CC   vulgaris, or for stimulating an immune response specific for a P. acnes
CC   protein. The polynucleotides can also be used as probes or primers for
CC   nucleic acid hybridisation. The vaccine composition is useful for the
CC   stimulation of an immune response against P. acnes, or for treating acne,
CC   and the kit is useful for performing a diagnostic assay. The present
CC   sequence represents a P. acnes DNA contig which is specifically claimed
CC   in the invention. Note: The sequence data for this patent did not form
CC   part of the printed specification, but was obtained in electronic format
CC   directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ   Sequence 1062 BP; 212 A; 338 C; 304 G; 205 T; 0 U; 3 Other;

  Query Match           100.0%;  Score 22;  DB 8;  Length 1062;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db         46 TCCTCCTGATATCTGCGCATTC 67


RESULT 9
ADZ67281/c
ID   ADZ67281 standard; DNA; 1135 BP.
XX
AC   ADZ67281;
XX
DT   30-JUN-2005  (first entry)
XX
DE   Frigoribacterium genus bacteria FERM P-19528 xylanase DNA.
XX
KW   xylanase; paper; pulp; ds.
XX
OS   Frigoribacterium.
XX
PN   JP2005102603-A.
XX
PD   21-APR-2005.
XX
PF   30-SEP-2003; 2003JP-00341110.
XX
PR   30-SEP-2003; 2003JP-00341110.
XX
PA   (DNIN ) DAINIPPON INK & CHEM INC.
PA   (UYNI-) UNIV NIPPON.
XX
DR   WPI; 2005-300063/31.
XX
PT   Novel xylanase capable of acting at preset pH, useful for processing pulp
PT   by degrading xylan in paper pulp at alkaline conditions.
XX
PS   Claim 6; SEQ ID NO 6; 15pp; Japanese.
XX
CC   The invention relates to a novel xylanase capable of acting at a pH
CC   ranging from 4-12. The invention further comprises: a Frigoribacterium
CC   genus bacteria capable of producing the novel xylanase; a
CC   Frigoribacterium genus bacteria sequence having a fully defined 1125 or
CC   1457 base pair sequence (ADZ67281 or ADZ67282) given in the specification
CC   and 95% or more homology to 16SrDNA; and a xylan processing agent for
CC   processing materials containing a polysaccharide of xylan, comprising the
CC   novel xylanase. The xylanase or xylan processing agent is useful for
CC   processing pulp. The Frigoribacterium genus bacteria is useful for
CC   producing the novel xylanase. The novel xylanase is useful for de-inking
CC   paper. The novel xylanase has excellent stability in alkaline conditions
CC   compared to conventional xylanase and enables efficient processing of
CC   paper pulp in a wide pH range (4-12). This polynucleotide sequence
CC   represents the Frigoribacterium genus bacteria FERM P-19528 xylanase DNA
CC   of the invention.
XX
SQ   Sequence 1135 BP; 285 A; 270 C; 359 G; 221 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 22;  DB 14;  Length 1135;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        324 TCCTCCTGATATCTGCGCATTC 303


RESULT 10
AAV24293/c
ID   AAV24293 standard; DNA; 1271 BP.
XX
AC   AAV24293;
XX
DT   14-SEP-1998  (first entry)
XX
DE   Mycobacterium tuberculosis 16S ribosomal RNA gene.
XX
KW   Antibacterial; antimycobacterial; oligonucleotide; infection; therapy;
KW   ribosome binding site; Shine-Dalgarno; ribosomal RNA; cystic firbosis;
KW   tuberculosis; ss.
XX
OS   Mycobacterium tuberculosis.
XX
PN   WO9814567-A2.
XX
PD   09-APR-1998.
XX
PF   30-SEP-1997;   97WO-US018094.
XX
PR   01-OCT-1996;   96US-0027729P.
XX
PA   (ADRE-) ADVANCED RES & TECHNOLOGY INST.
XX
PI   Martin WJ,  Wisniowski P;
XX
DR   WPI; 1998-240079/21.
XX
PT   Use of oligo:nucleotide(s) corresponding to bacterial 16S rRNA - for
PT   inhibiting bacterial protein expression and treating bacterial infection.
XX
PS   Claim 26; Page 60-61; 73pp; English.
XX
CC   This polynucleotide comprises the 16S ribosomal RNA (rRNA) gene of
CC   Mycobacterium tuberculosis. The invention relates to methods and
CC   compositions for the treatment of Gram-negative bacterial infections
CC   employing novel oligonucleotides as antimicrobial agents. The
CC   oligonucleotides are targeted to the Shine-Dalgarno (SD) region of
CC   prokaryotes to inhibit bacterial expression and hence inhibit bacterial
CC   infection. They preferably comprise 10-35 consecutive bases of the 3' end
CC   of a bacterial 16S rRNA (see also AAV24291-95). An oligonucleotide may
CC   also include a transport moiety and may have DNA phosphate modifications
CC   to increase nuclease resistance, or may be formulated in a liposome. A
CC   claimed method for treating a bacterial infection of a patient comprises
CC   administering a liposomal formulation of such an oligonucleotide. The
CC   oligonucleotides can be used particularly for treating bacterial
```

CC infections in pulmonary diseases such as cystic fibrosis or tuberculosis.
CC Since the SD sequence is not present in eukaryotic cells, the
CC oligonucleotides provide a pathogen-specific therapeutic method
XX
SQ Sequence 1271 BP; 260 A; 281 C; 430 G; 300 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 2; Length 1271;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 1088 TCCTCCTGATATCTGCGCATTC 1067

RESULT 11
AEA22410/c
ID AEA22410 standard; DNA; 1321 BP.
XX
AC AEA22410;
XX
DT 25-AUG-2005 (first entry)
XX
DE Mycobacterium kubicae 16S rRNA sequence SEQ ID NO:11.
XX
KW microorganism identification; 16S rRNA; 16S ribosomal RNA; ds.
XX
OS Mycobacterium kubicae.
XX
PN US2005130168-A1.
XX
PD 16-JUN-2005.
XX
PF 31-OCT-2003; 2003US-00697802.
XX
PR 31-OCT-2003; 2003US-00697802.
XX
PA (HANX/) HAN X.
PA (PHAM/) PHAM A S.
XX
PI Han X, Pham AS;
XX
DR WPI; 2005-424597/43.
XX
PT Determining a bacterium species comprises providing oligonucleotide
PT primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS Disclosure; SEQ ID NO 11; 74pp; English.
XX
CC The invention relates to a method (M1) for determining a bacterium
CC species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC extracting a genomic nucleotide from the bacterium to provide a
CC nucleotide template; (c) annealing a region of a nucleotide template to a
CC specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC complimentary fashion, the primer set designed to provide a product
CC having a predetermined size dictated by a complimentary primer set; (d)
CC amplifying the region of the nucleotide template to produce the product;
CC and (e) determining a species of a bacterium in a nucleotide sequence of
CC the product. Also described is an alternative method (M2) for determining
CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or
CC variations, or (iii) a first forward primer having consecutive bases of
CC an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC second reverse primer having consecutive bases of an UB-r of AEA22517-
CC AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC comparing the product from the specimen with a nucleotide sequence from a
CC database to determine the bacterium species present in the specimen. The
CC methods are useful for determining a bacterium species. The present
CC sequence represents a Mycobacterium kubicae 16S rRNA nucleotide sequence,
CC which is used in the exemplification of the present invention.
XX
SQ Sequence 1321 BP; 287 A; 314 C; 457 G; 263 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 1321;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 646 TCCTCCTGATATCTGCGCATTC 625

RESULT 12
AEA22400/c
ID AEA22400 standard; DNA; 1383 BP.
XX
AC AEA22400;
XX
DT 25-AUG-2005 (first entry)
XX
DE Mycobacterium abscessus 16S rRNA sequence SEQ ID NO:1.
XX
KW microorganism identification; 16S rRNA; 16S ribosomal RNA; ds.
XX
OS Mycobacterium abscessus.
XX
PN US2005130168-A1.
XX
PD 16-JUN-2005.
XX
PF 31-OCT-2003; 2003US-00697802.
XX
PR 31-OCT-2003; 2003US-00697802.
XX
PA (HANX/) HAN X.
PA (PHAM/) PHAM A S.
XX
PI Han X, Pham AS;
XX
DR WPI; 2005-424597/43.
XX
PT Determining a bacterium species comprises providing oligonucleotide
PT primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS Disclosure; SEQ ID NO 1; 74pp; English.
XX
CC The invention relates to a method (M1) for determining a bacterium
CC species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC extracting a genomic nucleotide from the bacterium to provide a
CC nucleotide template; (c) annealing a region of a nucleotide template to a
CC specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC complimentary fashion, the primer set designed to provide a product
CC having a predetermined size dictated by a complimentary primer set; (d)
CC amplifying the region of the nucleotide template to produce the product;
CC and (e) determining a species of a bacterium in a nucleotide sequence of
CC the product. Also described is an alternative method (M2) for determining
CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or

```
CC   variations, or (iii) a first forward primer having consecutive bases of
CC   an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC   second reverse primer having consecutive bases of an UB-r of AEA22517-
CC   AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC   comparing the product from the specimen with a nucleotide sequence from a
CC   database to determine the bacterium species present in the specimen. The
CC   methods are useful for determining a bacterium species. The present
CC   sequence represents a Mycobacterium abscessus 16S rRNA nucleotide
CC   sequence, which is used in the exemplification of the present invention.
XX
SQ   Sequence 1383 BP; 316 A; 328 C; 462 G; 277 T; 0 U; 0 Other;

  Query Match           100.0%;  Score 22;  DB 14;  Length 1383;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        636 TCCTCCTGATATCTGCGCATTC 615


RESULT 13
AAT45276/c
ID   AAT45276 standard; rRNA; 1391 BP.
XX
AC   AAT45276;
XX
DT   12-SEP-1997  (first entry)
XX
DE   Corynebacterium diphtheriae 16S rRNA.
XX
KW   Ribosomal RNA; species specific; detection; reverse transcription;
KW   primer; hybridisation probe; identification; ss.
XX
OS   Corynebacterium diphtheriae.
XX
FH   Key             Location/Qualifiers
FT   misc_feature    38. .59
FT                   /*tag= a
FT                   /note= "Defined as nucleotides 72-100"
FT   misc_feature    153. .170
FT                   /*tag= b
FT                   /note= "Defined as nucleotides 195-215"
FT   misc_feature    415. .431
FT                   /*tag= c
FT                   /note= "Defined as nucleotides 466-494"
FT   misc_feature    544. .567
FT                   /*tag= d
FT                   /note= "Defined as nucleotides 544-567"
FT   misc_feature    773. .787
FT                   /*tag= e
FT                   /note= "Defined as nucleotides 838-853"
FT   misc_feature    793. .808
FT                   /*tag= f
FT                   /note= "Defined as nucleotides 859-875"
FT   misc_feature    946. .965
FT                   /*tag= g
FT                   /note= "Defined as nucleotides 1013-1032"
XX
PN   FR2733755-A1.
XX
PD   08-NOV-1996.
XX
PF   03-MAY-1995;   95FR-00005494.
XX
PR   03-MAY-1995;   95FR-00005494.
XX
PA   (INMR ) BIO MERIEUX.
XX
PI   Mabilat C,  Ruimy R;
XX
DR   WPI; 1997-001708/01.
XX
PT   Fragments of Corynebacterium 16S RNA - useful as probes and primers for
PT   identifying Corynebacterium spp.
XX
PS   Claim 1; Fig 1; 60pp; French.
XX
CC   Fragments covering 90 % of the sequence of 16S ribosomal RNA were
CC   amplified from 28 strains of 25 different species of Corynebacterium by
CC   PCR using primers specific for eubacteria. The amplification products
CC   were sequenced and the sequences were aligned for comparison. It was
CC   found that certain regions, i.e. those corresponding to nucleotides 72-
CC   100, 195-215, 466-494, 608-631, 838-853, 859-875 and 1013-1033 in the 16S
CC   ribosomal RNA of C. diphtheriae (refer to features table for the present
CC   sequence), vary considerably between different species. Probes and
CC   primers comprising at least 5 nucleotides from one of these species-
CC   specific sequences, including the present sequence, or their complements,
CC   are useful to distinguish between different Corynebacterium species. DNA
CC   versions of the probes and primers are also included
XX
SQ   Sequence 1391 BP; 309 A; 317 C; 464 G; 1 T; 295 U; 5 Other;

  Query Match           100.0%;  Score 22;  DB 2;  Length 1391;
  Best Local Similarity  100.0%;  Pred. No. 1.1;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        661 TCCTCCTGATATCTGCGCATTC 640


RESULT 14
AEA22413/c
ID   AEA22413 standard; DNA; 1415 BP.
XX
AC   AEA22413;
XX
DT   25-AUG-2005  (first entry)
XX
DE   Mycobacterium paraffinicum 16S rRNA sequence SEQ ID NO:14.
XX
KW   microorganism identification; 16S rRNA; 16S ribosomal RNA; ds.
XX
OS   Mycobacterium paraffinicum.
XX
PN   US2005130168-A1.
XX
PD   16-JUN-2005.
XX
PF   31-OCT-2003; 2003US-00697802.
XX
PR   31-OCT-2003; 2003US-00697802.
XX
PA   (HANX/) HAN X.
PA   (PHAM/) PHAM A S.
XX
PI   Han X,  Pham AS;
XX
DR   WPI; 2005-424597/43.
XX
PT   Determining a bacterium species comprises providing oligonucleotide
PT   primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS   Disclosure; SEQ ID NO 14; 74pp; English.
XX
CC   The invention relates to a method (M1) for determining a bacterium
CC   species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC   extracting a genomic nucleotide from the bacterium to provide a
CC   nucleotide template; (c) annealing a region of a nucleotide template to a
CC   specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC   complimentary fashion, the primer set designed to provide a product
CC   having a predetermined size dictated by a complimentary primer set; (d)
CC   amplifying the region of the nucleotide template to produce the product;
CC   and (e) determining a species of a bacterium in a nucleotide sequence of
CC   the product. Also described is an alternative method (M2) for determining
```

CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or
CC variations, or (iii) a first forward primer having consecutive bases of
CC an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC second reverse primer having consecutive bases of an UB-r of AEA22517-
CC AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC comparing the product from the specimen with a nucleotide sequence from a
CC database to determine the bacterium species present in the specimen. The
CC methods are useful for determining a bacterium species. The present
CC sequence represents a Mycobacterium paraffinicum 16S rRNA nucleotide
CC sequence, which is used in the exemplification of the present invention.
XX
SQ Sequence 1415 BP; 307 A; 343 C; 480 G; 285 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 1415;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 656 TCCTCCTGATATCTGCGCATTC 635

RESULT 15
AEA22416/c
ID AEA22416 standard; DNA; 1416 BP.
XX
AC AEA22416;
XX
DT 25-AUG-2005 (first entry)
XX
DE Mycobacterium tuberculosis 16S rRNA sequence SEQ ID NO:17.
XX
KW microorganism identification; 16S rRNA; 16S ribosomal RNA; ds.
XX
OS Mycobacterium tuberculosis.
XX
PN US2005130168-A1.
XX
PD 16-JUN-2005.
XX
PF 31-OCT-2003; 2003US-00697802.
XX
PR 31-OCT-2003; 2003US-00697802.
XX
PA (HANX/) HAN X.
PA (PHAM/) PHAM A S.
XX
PI Han X, Pham AS;
XX
DR WPI; 2005-424597/43.
XX
PT Determining a bacterium species comprises providing oligonucleotide
PT primer set comprising SEQ-FOR and SEQ-REV in a complimentary fashion.
XX
PS Disclosure; SEQ ID NO 17; 74pp; English.
XX
CC The invention relates to a method (M1) for determining a bacterium
CC species. (M1) comprises: (a) culturing a bacterium from a specimen; (b)
CC extracting a genomic nucleotide from the bacterium to provide a
CC nucleotide template; (c) annealing a region of a nucleotide template to a
CC specific oligonucleotide primer set comprising SEQ-FOR and SEQ-REV in a
CC complimentary fashion, the primer set designed to provide a product
CC having a predetermined size dictated by a complimentary primer set; (d)
CC amplifying the region of the nucleotide template to produce the product;
CC and (e) determining a species of a bacterium in a nucleotide sequence of
CC the product. Also described is an alternative method (M2) for determining
CC a bacterium species comprising: (a) providing a specimen or a sample
CC having a template; (b) providing a pair of primers selected from: (i) a
CC first forward primer having consecutive bases of an AFB-f comprising any
CC of the 36 sequences of 15-22 bp (AEA22417-AEA22452), or their fragments
CC or variations and a first reverse primer having consecutive bases of an
CC AFB-r comprising any of the 36 sequences of 15-22 bp (AEA22453-AEA22488)
CC or their fragments or variations, (ii) a second forward primer having
CC consecutive bases of an UB-f comprising any of the 28 sequences of 15-21
CC bp (AEA22489-AEA22516) or their fragments or variations and a second
CC reverse primer having consecutive bases of an UB-r comprising any of the
CC 28 sequences of 15-21 bp (AEA22517-AEA22544) or their fragments or
CC variations, or (iii) a first forward primer having consecutive bases of
CC an AFB-f of AEA22417-AEA22452 or their fragments or variations and a
CC second reverse primer having consecutive bases of an UB-r of AEA22517-
CC AEA22544 or their fragments or variations; (c) the specimen; and (d)
CC comparing the product from the specimen with a nucleotide sequence from a
CC database to determine the bacterium species present in the specimen. The
CC methods are useful for determining a bacterium species. The present
CC sequence represents a Mycobacterium tuberculosis 16S rRNA nucleotide
CC sequence, which is used in the exemplification of the present invention.
XX
SQ Sequence 1416 BP; 309 A; 341 C; 481 G; 285 T; 0 U; 0 Other;

Query Match 100.0%; Score 22; DB 14; Length 1416;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 666 TCCTCCTGATATCTGCGCATTC 645

Search completed: April 7, 2006, 19:22:26
Job time : 222 secs

GenCore version 5.1.7

OM nucleic - nucleic search, using sw model

Run on: April 7, 2006, 19:15:09 ; Search time 1708.5 Seconds
(without alignments)
602.468 Million cell updates/sec

Title: US-10-697-802A-82
Perfect score: 22
Sequence: 1 tcctcctgatatctgcgcattc 22

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 41078325 seqs, 23393541228 residues

Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_htc:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_est7:*
9: gb_gss1:*
10: gb_gss2:*
11: gb_gss3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22	100.0	133	1	AW821632	AW821632 IL2-ST031
c	2	22	100.0	214	2	BE158593	BE158593 CM2-HT039
	3	22	100.0	269	3	BM194948	BM194948 L0703F05-
	4	22	100.0	269	5	BQ554727	BQ554727 H4029H12-
c	5	22	100.0	269	5	BQ554728	BQ554728 H4029H12-
c	6	22	100.0	537	3	BM130238	BM130238 pb28e03.y
c	7	22	100.0	725	7	CN204148	CN204148 Tor4539 G
c	8	21	95.5	382	6	CD164371	CD164371 ML1-0087T
c	9	21	95.5	645	8	DR884593	DR884593 JGI_CACX4
	10	21	95.5	722	6	CD164440	CD164440 ML1-0087T
c	11	21	95.5	744	6	CD164478	CD164478 ML1-0087T
	12	21	95.5	744	9	BZ781734	ii31b08.g
c	13	21	95.5	784	6	CB990869	CB990869 AGENCOURT
c	14	21	95.5	874	7	CO365133	CO365133 RTK1_23_G
	15	21	95.5	887	10	CL693661	CL693661 PRI0162a_
c	16	20.4	92.7	762	8	DR385950	DR385950 RTHG1_11_
c	17	20.4	92.7	817	7	CN207539	CN207539 Tor7952 G
c	18	20	90.9	403	9	BH740475	cpbav0005
	19	19.4	88.2	385	6	CD086973	CD086973 MC1-0033T
	20	19.4	88.2	506	8	DR072790	DR072790 RTDK1_28_
	21	19.4	88.2	517	8	DR072713	DR072713 RTDK1_28_
	22	19.4	88.2	617	6	CD096968	CD096968 ME1-0011T
c	23	19.4	88.2	740	6	CD164477	CD164477 ML1-0087T
	24	18.8	85.5	591	6	CD005515	CD005515 VVB052G08
	25	18.8	85.5	591	6	CD006477	CD006477 VVB086F03
c	26	18.8	85.5	677	10	CG898436	CG898436 pastbac07
	27	18.4	83.6	633	9	AQ655061	AQ655061 Sheared D
c	28	17.8	80.9	291	7	CO850294	CO850294 LM_SH5_00
c	29	17.8	80.9	309	5	BY354356	BY354356 BY354356
c	30	17.8	80.9	320	7	CO860316	CO860316 LM_SM5_00
c	31	17.8	80.9	339	6	CD087424	CD087424 MC1-0037T
	32	17.8	80.9	360	5	BX630302	BX630302 BX630302
c	33	17.8	80.9	388	8	CV842104	CV842104 ID0ADD3AH
c	34	17.8	80.9	431	2	BF469691	BF469691 UI-M-BH3-
c	35	17.8	80.9	443	8	CV848006	CV848006 ID0AEE4BE
c	36	17.8	80.9	446	3	BI499489	BI499489 ic58e06.y
	37	17.8	80.9	449	1	AI507902	AI507902 vp23g07.x
	38	17.8	80.9	474	9	AZ839667	AZ839667 2M0135G24
c	39	17.8	80.9	478	6	CD088114	CD088114 MC1-0048T
	40	17.8	80.9	480	5	BX566469	BX566469 BX566469
	41	17.8	80.9	510	6	CD086761	CD086761 MC1-0028U
c	42	17.8	80.9	514	3	BM130350	BM130350 pb29h10.y
	43	17.8	80.9	516	7	CK927912	CK927912 hasp010xn
c	44	17.8	80.9	517	3	BM130147	BM130147 pb27b07.y
c	45	17.8	80.9	525	3	BM130083	BM130083 pb26c05.y

ALIGNMENTS

RESULT 1
AW821632
LOCUS AW821632 133 bp mRNA linear EST 17-MAY-2000
DEFINITION IL2-ST0311-270300-059-E05 ST0311 Homo sapiens cDNA, mRNA sequence.
ACCESSION AW821632
VERSION AW821632.1 GI:7914626
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 133)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=IL2-ST0311-270300-059-E05&t3=2000-03-27&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 133.
FEATURES Location/Qualifiers
source 1. .133
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="ST0311"
/note="Organ: stomach; Vector: puc18; Site_1: SmaI;

Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 22; DB 1; Length 133;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 5 TCCTCCTGATATCTGCGCATTC 26

RESULT 2
BE158593/c
LOCUS BE158593 214 bp mRNA linear EST 21-JUN-2000
DEFINITION CM2-HT0393-301199-044-g05 HT0393 Homo sapiens cDNA, mRNA sequence.
ACCESSION BE158593
VERSION BE158593.1 GI:8621314
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 214)
AUTHORS Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Baia,G.S., Simpson,D.H., Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
PUBMED 10737800
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?t1=&t2=CM2-HT0393-301199-044-g05&t3=1999-11-30&t4=1)
Seq primer: puc 18 forward
High quality sequence start: 63
High quality sequence stop: 214.
FEATURES Location/Qualifiers
source 1. .214
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/dev_stage="Adult"
/clone_lib="HT0393"
/note="Organ: head_neck; Vector: puc18; Site_1: SmaI; Site_2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN

Query Match 100.0%; Score 22; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TCCTCCTGATATCTGCGCATTC 22
||||||||||||||||||||||
Db 192 TCCTCCTGATATCTGCGCATTC 171

RESULT 3
BM194948
LOCUS BM194948 269 bp mRNA linear EST 30-JAN-2002
DEFINITION L0703F05-3 NIA Mouse Germinal Center B Cell cDNA Library Mus musculus cDNA clone L0703F05 3', mRNA sequence.
ACCESSION BM194948
VERSION BM194948.1 GI:17746207
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 269)
AUTHORS Piao,Y., Kargul,G.J., Dudekula,D.B., Qian,Y., Lim,M.K., Klotz,E., Kelsoe,G., Hodes,R. and Ko,M.S.H.
TITLE Systematic Analyses of NIA Mouse Germinal Center B Cell cDNA Library
JOURNAL Unpublished (2001)
COMMENT Contact: Dawood B. Dudekula
Laboratory of Genetics
National Institute on Aging/National Institutes of Health
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA
Email: cdna@lgsun.grc.nia.nih.gov
Plate: L0703 row: F column: 05
Seq primer: -21M13 Forward
High quality sequence stop: 269
POLYA=Yes.
FEATURES Location/Qualifiers
source 1. .269
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="niaEST:L0703F05-3"
/db_xref="taxon:10090"
/clone="L0703F05"
/tissue_type="Germinal Center B Cell"
/lab_host="DH10B"
/clone_lib="NIA Mouse Germinal Center B Cell cDNA Library"
/note="Vector: pSPORT1 (Invitrogen); Site_1: SalI; Site_2: NotI; Mouse cDNA project by the Laboratory of Genetics, National Institute on Aging (NIA), Intramural Research Program, NIH (http://lgsun.grc.nia.nih.gov/cDNA).FACS-sorted Germinal Center B cells were provided by Drs. Richard Hodes, Emily Klotz (National Institute on Aging and National Cancer Institute, USA) and Garnett Kelsoe (Duke University, USA). Double-stranded cDNAs were synthesized from 0.46 ug of total RNA with an Oligo(dT) primer [Invitrogen: 5'-pGACTAGTTCTAGATCGCGAGCGGCCGCCCTTTTTTTTTTTTTT-3'], treated with T4 DNA polymerase, and purified by ethanol-precipitation. The cDNAs were ligated to Lone-linker LL-Sal3 (Ref. Development 127: 1737-1749 (2000) [PMID: 10725249]), purified by phenol/chloroform, and separated from free linkers by Centricon 100. Then, cDNAs were amplified by long-range high fidelity PCR using Ex Taq polymerase (Takara) and purified by phenol/chloroform, followed by Centricon 100 purification. The cDNAs were digested with SalI and NotI enzymes and cloned into SalI/NotI site of pSPORT1 plasmid vector. The DH10B E. coli host was transformed with the ligation mixture by the standard chemical method. The average insert size is 1.2 kb. The library was constructed by Yulan Piao (NIA)."

ORIGIN

Query Match 100.0%; Score 22; DB 3; Length 269;

```
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches    22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        143 TCCTCCTGATATCTGCGCATTC 164


RESULT 4
BQ554727
LOCUS       BQ554727                269 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  H4029H12-3 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
            H4029H12 3', mRNA sequence.
ACCESSION   BQ554727
VERSION     BQ554727.1  GI:21455615
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
            Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T.,
            Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
  TITLE     Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
  JOURNAL   Genome Res. 12 (12), 1999-2003 (2002)
   PUBMED   12466305
COMMENT     Other_ESTs: H4029H12-5
            Contact: Yong Qian
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
            Plate: H4029  row: H  column: 12
            Seq primer: -21M13 Forward
            High quality sequence stop: 269
            POLYA=Yes.
FEATURES             Location/Qualifiers
     source          1. .269
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="niaEST:H4029H12-3"
                     /db_xref="taxon:10090"
                     /clone="H4029H12"
                     /sex="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 7,407 clones from more
                     than 20 cDNA libraries."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 5;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches    22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        143 TCCTCCTGATATCTGCGCATTC 164


RESULT 5
BQ554728/c
LOCUS       BQ554728                269 bp    mRNA    linear   EST 20-JUN-2002
DEFINITION  H4029H12-5 NIA Mouse 7.4K cDNA Clone Set Mus musculus cDNA clone
            H4029H12 5', mRNA sequence.
ACCESSION   BQ554728
VERSION     BQ554728.1  GI:21455616
KEYWORDS    EST.
SOURCE      Mus musculus (house mouse)
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 269)
  AUTHORS   VanBuren,V., Piao,Y., Dudekula,D.B., Qian,Y., Carter,M.G.,
            Martin,P.R., Stagg,C.A., Bassey,U., Aiba,K., Hamatani,T.,
            Kargul,G.J., Luo,A.G., Kelso,J., Hide,W. and Ko,M.S.H.
  TITLE     Assembly, verification, and initial annotation of NIA 7.4K mouse
            cDNA clone set
  JOURNAL   Genome Res. 12 (12), 1999-2003 (2002)
   PUBMED   12466305
COMMENT     Other_ESTs: H4029H12-3
            Contact: Yong Qian
            Laboratory of Genetics
            National Institute on Aging/National Institutes of Health
            333 Cassell Drive, Suite 3000, Baltimore, MD 21224-6820, USA
            Email: cdna@lgsun.grc.nia.nih.gov
            This clone set has been freely distributed to the community. Please
            visit http://lgsun.grc.nia.nih.gov/cDNA/NIA_7_4k.html for details.
            Plate: H4029  row: H  column: 12
            Seq primer: -21M13 Reverse
            High quality sequence stop: 269
            POLYA=No.
FEATURES             Location/Qualifiers
     source          1. .269
                     /organism="Mus musculus"
                     /mol_type="mRNA"
                     /strain="C57BL/6"
                     /db_xref="niaEST:H4029H12-5"
                     /db_xref="taxon:10090"
                     /clone="H4029H12"
                     /sex="mixed"
                     /dev_stage="mixed"
                     /lab_host="DH10B"
                     /clone_lib="NIA Mouse 7.4K cDNA Clone Set"
                     /note="Vector: pSPORT1; Site_1: SalI; Site_2: NotI; This
                     clone is among a rearrayed set of 7,407 clones from more
                     than 20 cDNA libraries."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 5;  Length 269;
  Best Local Similarity   100.0%;  Pred. No. 3.9;
  Matches    22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        127 TCCTCCTGATATCTGCGCATTC 106


RESULT 6
BM130238/c
LOCUS       BM130238                537 bp    mRNA    linear   EST 27-NOV-2001
DEFINITION  pb28e03.y1 Anc caninum L3 serum stim pAMP1 v1 Chiapelli McCarter
            Ancylostoma caninum cDNA 5', mRNA sequence.
ACCESSION   BM130238
VERSION     BM130238.1  GI:17124790
KEYWORDS    EST.
SOURCE      Ancylostoma caninum (dog hookworm)
  ORGANISM  Ancylostoma caninum
            Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Strongylida;
            Ancylostomatoidea; Ancylostomatidae; Ancylostomatinae; Ancylostoma.
REFERENCE   1  (bases 1 to 537)
  AUTHORS   McCarter,J., Clifton,S., Chiapelli,B., Pape,D., Martin,J.,
            Wylie,T., Dante,M., Marra,M., Hillier,L., Kucaba,T., Theising,B.,
            Bowers,Y., Gibbons,M., Ritter,E., Bennett,J., Franklin,C.,
            Tsagareishvili,R., Ronko,I., Kennedy,S., Maguire,L., Beck,C.,
            Underwood,K., Steptoe,M., Allen,M., Person,B., Swaller,T.,
            Harvey,N., Schurk,R., Kohn,S., Shin,T., Jackson,Y., Cardenas,M.,
            McCann,R., Waterston,R. and Wilson,R.
```

```
  TITLE     The Washington Univ. Nematode EST Project, 1999
  JOURNAL   Unpublished (1999)
COMMENT     Contact: McCarter JP
            The Washington Univ. Nematode EST Project, 1999
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            The library was constructed by Brandi Chiapelli and Dr. James
            McCarter (bchiapel@watson.wustl.edu & jmccarte@watson.wustl.edu) at
            Washington University, St. Louis. DNA Sequencing by: Washington
            University Genome Sequencing Center St. Louis. Nematodes were
            provided by Dr. Prema Arasu of North Carolina State University.
            High quality sequence stop: 395.
FEATURES             Location/Qualifiers
     source          1. .537
                     /organism="Ancylostoma caninum"
                     /mol_type="mRNA"
                     /db_xref="taxon:29170"
                     /dev_stage="serum stimulated L3"
                     /lab_host="DH10B"
                     /clone_lib="Anc caninum L3 serum stim pAMP1 v1 Chiapelli
                     McCarter"
                     /note="Vector: pAMP1 (Gibco); Site_1: NotI; Site_2: SalI;
                     The library was constructed by Brandi Chiapelli and Dr.
                     James  McCarter at Washington University, St. Louis.  The
                     cDNA was made by using Dynabead oligo-dT priming (Dynal).
                     PCR based library using a modified  protocol from the
                     SMART PCR cDNA Synthesis Kit from Clontech.  Directionally
                     cloned into the UDG sites of pAMP1.  Nematodes were
                     provided by Dr. Prema Arasu of North Carolina State
                     University."
ORIGIN

  Query Match             100.0%;  Score 22;  DB 3;  Length 537;
  Best Local Similarity   100.0%;  Pred. No. 4.4;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        151 TCCTCCTGATATCTGCGCATTC 130


RESULT 7
CN204148/c
LOCUS       CN204148                725 bp    mRNA    linear   EST 30-APR-2004
DEFINITION  Tor4539 Gametophyte rehydration Library Tortula ruralis cDNA, mRNA
            sequence.
ACCESSION   CN204148
VERSION     CN204148.1  GI:46900879
KEYWORDS    EST.
SOURCE      Tortula ruralis
  ORGANISM  Tortula ruralis
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
            Bryopsida; Dicranidae; Pottiales; Pottiaceae; Tortula.
REFERENCE   1  (bases 1 to 725)
  AUTHORS   Oliver,M.J., Dowd,S.E., Zaragosa,J., Mauget,S.A. and Payton,P.R.
  TITLE     The rehydration transcriptome of the desiccation-tolerant bryophyte
            Tortula ruralis: transcript classification and analysis
  JOURNAL   BMC Genomics 5 (1), 89 (2004)
   PUBMED   15546486
COMMENT     Contact: Oliver Melvin J
            Plant Stress Lab
            USDA-ARS
            3810 4th St, Lubbock, TX 79415, USA
            Tel: 806-749-5560
            Fax: 806-723-5272
            Email: moliver@lbk.ars.usda.gov
            PCR PRimers
            FORWARD: GTTTTCCCAGTCACGAC
            BACKWARD: CAGGAAACAGCTATGAC.
FEATURES             Location/Qualifiers
     source          1. .725
                     /organism="Tortula ruralis"
                     /mol_type="mRNA"
                     /db_xref="taxon:38588"
                     /clone_lib="Gametophyte rehydration Library"
                     /note="Organ: Green Gametophyte; Vector: pSport1; Site_1:
                     SalI; Site_2: NotI"
ORIGIN

  Query Match             100.0%;  Score 22;  DB 7;  Length 725;
  Best Local Similarity   100.0%;  Pred. No. 4.7;
  Matches   22;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATTC 22
              ||||||||||||||||||||||
Db        686 TCCTCCTGATATCTGCGCATTC 665


RESULT 8
CD164371/c
LOCUS       CD164371                382 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ML1-0087T-R218-E03-U.G ML1-0087 Schistosoma mansoni cDNA clone
            ML1-0087T-R218-E03.G, mRNA sequence.
ACCESSION   CD164371
VERSION     CD164371.1  GI:34701042
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 382)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
   PUBMED   12973350
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ML1-0087T-R218  row: 3  column: E.
FEATURES             Location/Qualifiers
     source          1. .382
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="ML1-0087T-R218-E03.G"
                     /sex="mixed pool"
                     /dev_stage="miracidium"
                     /clone_lib="ML1-0087"
                     /note="Vector: pGEM T-easy"
ORIGIN

  Query Match             95.5%;  Score 21;  DB 6;  Length 382;
  Best Local Similarity   100.0%;  Pred. No. 13;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;
```

```
Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        226 TCCTCCTGATATCTGCGCATT 206


RESULT 9
DR884593/c
LOCUS       DR884593                645 bp    mRNA    linear   EST 01-AUG-2005
DEFINITION  JGI_CACX492.fwd NIH_XGC_tropMet5 Xenopus tropicalis cDNA clone
            IMAGE:7796789 5', mRNA sequence.
ACCESSION   DR884593
VERSION     DR884593.1  GI:71573845
KEYWORDS    EST.
SOURCE      Xenopus tropicalis (western clawed frog)
  ORGANISM  Xenopus tropicalis
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidea; Pipidae;
            Xenopodinae; Xenopus; Silurana.
REFERENCE   1  (bases 1 to 645)
  AUTHORS   Richardson,P., Lucas,S., Rokhsar,D., Detter,J.C., Ng,D.C.,
            Brokstein,P. and Lindquist,E.A.
  TITLE     DOE Joint Genome Institute Xenopus tropicalis EST project
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Lindquist,E.A., Richardson,P.
            DOE Joint Genome Institute
            2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            Tel: 925 296 5600
            Fax: 925 296 5710
            Email: cdna@jgi-psf.org
            Tissue Procurement: Dan Buchholz (Yun-Bo Shi Laboratory, NIH)
            cDNA Library Preparation: DOE Joint Genome Institute:
            http://www.jgi.doe.gov
            DNA Sequencing: DOE Joint Genome Institute: http://www.jgi.doe.gov
            Clone Distribution: I.M.A.G.E. Consortium/LLNL:
            http://image.llnl.gov
            Naming Conventions: EST name is generated by the concatenation of
            the JGI Clone Id and the direction of sequencing. The suffix '.fwd'
            indicates a forward sequencing read of the insert. It does not
            necessarily reflect the orientation of the insert.
            Small Insert: Based upon one or more sequencing reads of this clone
            where vector sequence was present at both ends, this clone has been
            determined to contain a cDNA insert on the order of 600-1000 bases.
            Plate: CACX 0005  row: h  column: 3
            High quality sequence stop: 616.
FEATURES             Location/Qualifiers
     source          1. .645
                     /organism="Xenopus tropicalis"
                     /mol_type="mRNA"
                     /db_xref="taxon:8364"
                     /clone="IMAGE:7796789"
                     /tissue_type="whole embryo"
                     /dev_stage="Metamorphic (st.62)"
                     /lab_host="ElectroMAX DH10B"
                     /clone_lib="NIH_XGC_tropMet5"
                     /note="Vector: pCMVsport6; Site_1: SalI; Site_2: NotI;
                     This library was made from dT primed cDNA and cloned into
                     Invitrogen pCMVsport6 vector. The work was done at DOE
                     Joint Genome Institute. Poly A RNA were primed with 5'
                     GACTAGTTCTAGATCGCGAG CGGCCGCCCTTTTTTTTTTTTTTT 3'. CDNA
                     were ligated to SalI adapter (5' TCGACCCACGCGTCCG and
                     5'CGGACGCGTGGG), digested with NotI, size fractionated in
                     1.1% agarose gel electrophoresis and ligated into NotI and
                     SalI digested pCMVsport6 vector."
ORIGIN

  Query Match             95.5%;  Score 21;  DB 8;  Length 645;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        426 TCCTCCTGATATCTGCGCATT 406


RESULT 10
CD164440
LOCUS       CD164440                722 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ML1-0087T-R244-G10-U.G ML1-0087 Schistosoma mansoni cDNA clone
            ML1-0087T-R244-G10.G, mRNA sequence.
ACCESSION   CD164440
VERSION     CD164440.1  GI:34701106
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
REFERENCE   1  (bases 1 to 722)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
   PUBMED   12973350
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ML1-0087T-R244  row: 10  column: G.
FEATURES             Location/Qualifiers
     source          1. .722
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="ML1-0087T-R244-G10.G"
                     /sex="mixed pool"
                     /dev_stage="miracidium"
                     /clone_lib="ML1-0087"
                     /note="Vector: pGEM T-easy"
ORIGIN

  Query Match             95.5%;  Score 21;  DB 6;  Length 722;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db         26 TCCTCCTGATATCTGCGCATT 46


RESULT 11
CD164478/c
LOCUS       CD164478                744 bp    mRNA    linear   EST 14-SEP-2003
DEFINITION  ML1-0087T-R250-E07-U.G ML1-0087 Schistosoma mansoni cDNA clone
            ML1-0087T-R250-E07.G, mRNA sequence.
ACCESSION   CD164478
VERSION     CD164478.1  GI:34701143
KEYWORDS    EST.
SOURCE      Schistosoma mansoni
  ORGANISM  Schistosoma mansoni
            Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea;
            Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.
```

```
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Verjovski-Almeida,S., DeMarco,R., Martins,E.A.L., Guimaraes,P.E.M.,
            Ojopi,E.P.B., Paquola,A.C.M., Piazza,J.P., Nishiyama,M.Y. Jr.,
            Kitajima,J.P., Adamson,R.E., Ashton,P.D., Bonaldo,M.F.,
            Coulson,P.S., Dillon,G.P., Farias,L.P., Gregorio,S.P., Ho,P.L.,
            Leite,R.A., Malaquias,L.C.C., Marques,R.C.P., Miyasato,P.A.,
            Nascimento,A.L.T.O., Ohlweiler,F.P., Reis,E.M., Ribeiro,M.A.,
            Sa,R.G., Stukart,G.C., Soares,M.B., Gargioni,C., Kawano,T.,
            Rodrigues,V., Madeira,A.M.B.N., Wilson,R.A., Menck,C.F.M.,
            Setubal,J.C., Leite,L.C.C. and Dias-Neto,E.
  TITLE     Transcriptome analysis of the acoelomate human parasite Schistosoma
            mansoni
  JOURNAL   Nat. Genet. 35 (2), 148-157 (2003)
   PUBMED   12973350
COMMENT     Contact: Dr. Sergio Verjovski-Almeida
            Departamento de Bioquimica
            Instituto de Quimica - Universidade de Sao Paulo
            Av. Prof. Lineu Prestes 748 sala 1200, 05508-900 Sao Paulo - SP,
            Brasil
            Tel: +55-11-3091-2173
            Fax: +55-11-3091-2186
            Email: verjo@iq.usp.br
            This sequence was derived from the FAPESP Schistosoma mansoni EST
            Genome Project. All sequences in the project were assembled and
            annotated. This entry and all the assembled sequences can be seen
            in the following URL http://bioinfo.iq.usp.br/schisto/
            Plate: ML1-0087T-R250  row: 7  column: E.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Schistosoma mansoni"
                     /mol_type="mRNA"
                     /db_xref="taxon:6183"
                     /clone="ML1-0087T-R250-E07.G"
                     /sex="mixed pool"
                     /dev_stage="miracidium"
                     /clone_lib="ML1-0087"
                     /note="Vector: pGEM T-easy"
ORIGIN

  Query Match             95.5%;  Score 21;  DB 6;  Length 744;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        715 TCCTCCTGATATCTGCGCATT 695


RESULT 12
BZ781734
LOCUS       BZ781734                744 bp    DNA     linear   GSS 14-MAR-2003
DEFINITION  ii31b08.g1 WGS-SbicolorF (DH5a methyl filtered) Sorghum bicolor
            genomic clone ii31b08; genomic survey sequence.
ACCESSION   BZ781734
VERSION     BZ781734.1  GI:28960179
KEYWORDS    GSS.
SOURCE      Sorghum bicolor (sorghum)
  ORGANISM  Sorghum bicolor
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Sorghum.
REFERENCE   1  (bases 1 to 744)
  AUTHORS   Rabinowicz,P.D., O'Shaughnessy,A.L., Balija,V., Dedhia,N.,
            Katzenburger,F., King,L., Miller,B., Muller,S., Nascimento,L.,
            Zutavern,T., Palmer,L., McCombie,W.R. and Martienssen,R.A.
  TITLE     Genomic shotgun sequences from Sorghum bicolor (methyl-filtered)
  JOURNAL   Unpublished (2002)
COMMENT     Contact: W. Richard McCombie
            Lita Annenberg Hazen Genome Sequencing Center
            Cold Spring Harbor Laboratory
            PO Box 100, Cold Spring Harbor, NY 11724, USA
            Tel: 516 367 8884
            Fax: 516 367 8874
            Email: mccombie@cshl.org
            Plate: ii31  row: b  column: 08
            Seq primer: -21M13UnivRev
            Class: shotgun
            High quality sequence stop: 744.
FEATURES             Location/Qualifiers
     source          1. .744
                     /organism="Sorghum bicolor"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:4558"
                     /clone="ii31b08"
                     /lab_host="DH5a"
                     /clone_lib="WGS-SbicolorF (DH5a methyl filtered)"
                     /note="Site_1: Xba I; Site_2: Xba I; The vector was
                     digested with XbaI and one nucleotide was added by fill in
                     in the recessive 3' end. The genomic DNA was nebulized,
                     end repaired, adaptor ligated and size fractionated using
                     sephadex. The resulting fragments were between 0.8 and 3
                     kb and were cloned into the vector (.x/y reads in M13mp19,
                     .b/g reads in pUC19). The same ligation was transformed
                     into DH5a."
ORIGIN

  Query Match             95.5%;  Score 21;  DB 9;  Length 744;
  Best Local Similarity   100.0%;  Pred. No. 15;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        672 TCCTCCTGATATCTGCGCATT 692


RESULT 13
CB990869/c
LOCUS       CB990869                784 bp    mRNA    linear   EST 01-MAY-2003
DEFINITION  AGENCOURT_13620403 NIH_MGC_148 Homo sapiens cDNA clone
            IMAGE:30338309 5', mRNA sequence.
ACCESSION   CB990869
VERSION     CB990869.1  GI:30285389
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1  (bases 1 to 784)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Dr. Stefan Hansson
             cDNA Library Preparation: Michael J. Brownstein (NHGRI) with help
            and advice from Piero Carninci (RIKEN)
             cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: Agencourt Bioscience Corporation
             Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: NDAM364  row: m  column: 06
            High quality sequence stop: 539.
FEATURES             Location/Qualifiers
     source          1. .784
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:30338309"
                     /tissue_type="pre-eclamptic placenta"
                     /lab_host="DH10B TonA"
                     /clone_lib="NIH_MGC_148"
                     /note="Organ: placenta; Vector: pBluescriptR; Site_1:
                     aII-XhoI; Site_2: BamH; Library is oligo-dT primed and
                     directionally cloned using primer
```

5'-TTTTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.3 kb and normalized to ROT 5. This is a primary library enriched for full-lenght clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 95.5%; Score 21; DB 6; Length 784;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        595 TCCTCCTGATATCTGCGCATT 575
```

RESULT 14
CO365133/c
LOCUS CO365133 874 bp mRNA linear EST 29-JUN-2004
DEFINITION RTK1_23_G09.g1_A029 Roots minus potassium Pinus taeda cDNA clone RTK1_23_G09_A029 5', mRNA sequence.
ACCESSION CO365133
VERSION CO365133.1 GI:49446450
KEYWORDS EST.
SOURCE Pinus taeda (loblolly pine)
ORGANISM Pinus taeda
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
REFERENCE 1 (bases 1 to 874)
AUTHORS Pratt,L., Cordonnier-Pratt,M.-M., Lorenz,W.W., Zimmermann,C. and Dean,J.F.D.
TITLE An EST database from potassium-deficient loblolly pine (Pinus taeda) roots
JOURNAL Unpublished (2004)
COMMENT Other_ESTs: RTK1_23_G09.b1_A029
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
RNA prepared and library constructed by W. Walter Lorenz (School of Forest Resources, University of Georgia); plant material prepared by Craig Zimmermann (School of Forest Resources, University of Georgia) using rooted cuttings provided by the Forest Biology Research Cooperative (FBRC) and the CCLONES project a the University of Florida; sequencing done in the Laboratory for Genomics and Bioinformatics, University of Georgia. Sequence ends have been trimmed to exclude vector and regions below Phred quality 16. Three-prime sequences are presented as their reverse complement and have been trimmed to exclude polyA.
Seq primer: JENREV (CAGGAAACAGCTATGACC).
FEATURES Location/Qualifiers
source 1. .874
/organism="Pinus taeda"
/mol_type="mRNA"
/strain="3 CCLONES"
/db_xref="taxon:3352"
/clone="RTK1_23_G09_A029"
/lab_host="DH10B-T1 phage-resistant E. coli"
/clone_lib="Roots minus potassium"
/note="Organ: Root; Vector: pSL1180; Site_1: EcoRI; Site_2: XhoI; The library was prepared from polyA+ RNA from the roots of 1-year-old loblolly pine (Pinus taeda) cuttings that were rooted and then planted in washed sand. The rooted cuttings were maintained for 117 days (July 2003 harvest) under ambient conditions in a local greenhouse. They were kept on a weekly regimen of 0.5x nutrient-complete Hoagland's solution and supplemented with additional water sufficient to maintain a 15% soil moisture content. For twenty-eight days (28 d) prior to harvesting roots for mRNA preparation, the trees received Hoagland's solution lacking potassium (K) to induce a potassium-deficiency. Double-stranded cDNA was cloned unidirectionally into pSL1180. Inserts can be excised with EcoRI (5' end) and XhoI (3' end)."

ORIGIN

Query Match 95.5%; Score 21; DB 7; Length 874;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        651 TCCTCCTGATATCTGCGCATT 631
```

RESULT 15
CL693661
LOCUS CL693661 887 bp DNA linear GSS 10-JUL-2004
DEFINITION PRI0162a_H01_2 - PRI0162a.BR (887) Mixed stage fosmid library of P. pacificus var. California Pristionchus pacificus genomic, genomic survey sequence.
ACCESSION CL693661
VERSION CL693661.1 GI:50215569
KEYWORDS GSS.
SOURCE Pristionchus pacificus
ORGANISM Pristionchus pacificus
Eukaryota; Metazoa; Nematoda; Chromadorea; Diplogasterida; Neodiplogasteridae; Pristionchus.
REFERENCE 1 (bases 1 to 887)
AUTHORS Srinivasan,J., Otto,G.W., Kahlow,U., Geisler,R. and Sommer,R.J.
TITLE AppaDB: an AcedB database for the nematode satellite organism Pristionchus pacificus
JOURNAL Nucleic Acids Res. 32 (1), D421-D422 (2004)
PUBMED 14681447
COMMENT Contact: Sommer RJ
Evolutionary Biology
Max-Planck-Institute for Developmental Biology
Spemannstr. 37-39, Tuebingen D-72076, Germany
Tel: 00497071601371
Fax: 00497071601498
Email: ralf.sommer@tuebingen.mpg.de
This library was generated at Caltech, Pasadena, USA and end sequenced at Vancouver, Canada.
Seq primer: T7
Class: fosmid ends.
FEATURES Location/Qualifiers
source 1. .887
/organism="Pristionchus pacificus"
/mol_type="genomic DNA"
/strain="California"
/db_xref="taxon:54126"
/clone_lib="Mixed stage fosmid library of P. pacificus var. California"
/note="Vector: pEpifos-5 Fosmid vector"

ORIGIN

Query Match 95.5%; Score 21; DB 10; Length 887;
Best Local Similarity 100.0%; Pred. No. 16;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy          1 TCCTCCTGATATCTGCGCATT 21
              |||||||||||||||||||||
Db        210 TCCTCCTGATATCTGCGCATT 230
```

Search completed: April 7, 2006, 20:19:41
Job time : 1715.5 secs